

The Vatican and embryology

Last week's Instruction to Roman Catholics on in vitro fertilization and related matters is based on an unrealistic and insensitive appreciation of what modern science is like.

THE Vatican, represented by the Congregation of the Faith, has done itself more harm than good with its statement last week on the lawful and illicit uses of new techniques in human embryology. While claiming a foundation in modern science, the statement emerges as a restrictive prescription out of tune with the temper of science as it is. But natural scoffers should pause, at least for long enough to read the document put out from Rome, before denouncing it as irrelevant. At the least, the document could further confuse several important issues of public policy. If taken at its face value it could be a drag on research and the therapeutic techniques that will follow.

Among the several recommendations in what is technically an Instruction to members of the Church, but which is also intended as a legislative programme for sympathetic civil governments, the most striking is the interdiction of the use of two spouses' own gametes as the starting-point for *in vitro* fertilization (IVF). Everybody knows that Rome is against contraception and abortion, but why, people are asking, should two mutually faithful people who happen to be mutually infertile be deprived of children when there is a technique that may give them relief from their frustration? As may be expected, the Vatican's argument is subtle, consistent and consonant with previous declarations on related matters. Its defect is an oversimple view of science, which is nevertheless offered as the foundation of last week's Instruction.

The new Instruction is derived, with only a little contemporary tuning, from two moral principles — that the lives of individual human beings should be respected absolutely and that there must be an unbreakable and reciprocal connection between all sexual acts and the process of procreation. The first principle is nowhere disputed, although the Vatican characteristically goes further than would most (but not all) biologists in its insistence that life begins with the union of two haploid gametes to form a diploid zygote. The second principle, on the other hand, is distinctively Roman Catholic in its strict application, and is the origin of the Church's condemnation of masturbation and contraception, both practices lacking what the Instruction describes as "openness to procreation".

Science affirmed

For the past several years, one of the dilemmas forced on the Vatican by developments in biology has been the question whether the second principle applies in reverse, whether procreation without natural conjugal intercourse should be discouraged or even disallowed. The decision, that it must be forbidden, while amply supported in the Instruction with references to recent and past doctrinal authorities is where the trouble starts.

Some time ago, the Vatican demanded to know of Galileo whether it was his opinion that the Sun goes round the Earth or the other way about. The legend is that the old man chose to tell what he believed to be a lie even though, in these relativistic days, he would have been well within his rights to answer "either or neither, as you wish". The Church has learned the lesson of that sad encounter and is now ready to embrace scientific discovery even if, on the evidence of the Instruction, selectively

and even perversely. Unfortunately, the conclusion that the two-way application of the doctrine on conjugal intercourse and procreation is, even in the Church's own terms, both unnecessary and a hostage to the fortune of future scientific discovery. In reaching its conclusion, the Vatican has relied on too literal an appreciation of what science is about. It sets too much store by particular discoveries and too little by the objective of the enterprise, which is understanding.

This is nicely illustrated by the issue of deciding when, in the course of human development, life begins. Last week's Instruction takes as its starting-point "recent findings of human biological science" that all zygotes are genetically distinctive, so that the "biological identity of a new human individual is already constituted" once fertilization is complete. Mindful, no doubt, of the old arguments about when the soul may be considered to enter the body, the Instruction admits that there is "no experimental datum" to help decide, but asks the rhetorical question "how could a human individual not be a human person?".

Life's beginning

The case for banning even "homologous" IVF, where the parents provide the gametes, is built on this foundation. The first principle of the absolute sanctity of life after fertilization provides an elementary justification: IVF at present entails the production of 'spare' embryos that cannot be discarded without breaking the Church's law. The Instruction, anticipating that this objection may in future be met by the improvement of clinical technique, goes on to say that IVF would even then remain illicit because of the principle of the indissoluble linkage of conjugal intercourse and procreation. First, it says, because IVF entrusts "the life and identity" of the embryo into the "power of physicians and biologists", establishing the "domination of technology over the origin and the destiny of the human person", it offends intolerably against that person's dignity. Second, and in any case, because IVF is by definition different from conjugal intercourse, children emerging from the process are deprived of "proper perfection". Tolerantly, however, the Instruction reminds members of the Church that even children unfortunate enough to be conceived in this way "must be brought up with love" and couples who happen to be infertile that marriage does not bring "the right to have a child" but only the right "to perform those natural acts which are intrinsically directed towards procreation".

These conclusions will mystify most who are not Roman Catholics (and disappoint many who are). The remark that each embryo's DNA is distinctive is indisputable, but does not lead logically to the conclusion that each embryo is a "new human individual"; for the time being, it seems that the hospitality of a uterus, most probably a human uterus, is required to accomplish that. (The Instruction also bans trans-specific gestation and the artificial uterus.) If further investigation should show, which is not improbable, that the sojourn of an embryo in a uterus is somehow indispensable to its maturation, would the whole set of interdictions promulgated last week then collapse like a house of cards? Of course not, but for reasons quite different from what the Instruction calls the "cold logic" of its prescriptions.



Ironically, while the Congregation of the Faith has abandoned the old quest for a knowledge of when the soul enters the body, others have been moving in the opposite direction, but from different motives. In laboratories, for example, it is important to know when the development of artificially fertilized embryos must be stopped for fear of accumulating unwanted and inviable human individuals in Petri dishes. For want of a sufficient knowledge of embryogenesis, responsible civil jurisdictions have settled for time-limits acknowledged to be somewhat arbitrary, such as the 14-day limit proposed three years ago in Britain by the Warnock committee. Last week's Instruction, which also forbids research with, as distinct from therapy on, living embryos would enormously impede the search for a more viable understanding of where the limits should be drawn so as not to offend those who work in laboratories, and who are no readier than the next person to throw living "human individuals" away simply because they can be seen only with a microscope. In its preamble, the document asserts that science cannot be held to be "morally neutral", and insists that "science without conscience" is offensive. Would it not have been charitable to acknowledge that researchers have been among the most active in the search for a seemingly definition of when life begins?

That would also have been prudent. Over the years, the Church has accumulated a great burden for itself and its members by its position on the absolute sanctity of human life. Morally and even intellectually, the position deserves respect. But what if it is unnecessary? What if it is equally respectable, morally and intellectually, to found a view of when life begins on an account of early embryogenesis more rounded than that now accessible based on an appreciation of the development of the early nervous system in the three weeks after fertilization, on the emergence of distinctively mammalian characteristics in the days that follow or on other observations yet to be carried out? In these post-galilean times, the only certainty is that there will be no simple way of drawing a distinction between the living and the human, but outsiders cannot but ask whether the Church itself would not benefit from a successful search towards that end. Then it need no longer be impaled on the hooks of its policy on contraception and, now, on IVF.

Cruel doctrine

Much the same is true of what the Instruction says of the link between conjugal intercourse and procreation. There is a body of opinion much wider than the Church that the stability of the family should be an objective of public policy. In many jurisdictions, the principle is accepted that there is no inalienable right of couples "to have" children. But can it seriously be held, in the face of evidence from social scientists and clinical psychiatrists, that the moral error in the relief of the frustrations of sterile marriages outweighs the evident benefits of physicians' admittedly artificial intervention? This is a cruel doctrine, as the Instruction admits. Is it really necessary, and can it be wise?

The Vatican, of course, will say that prudence should be no part of its calculation, only moral rectitude. By the Church's own lights, nobody can quarrel with that position. The issue is simply whether the experimental basis of its moral position is as secure as the Congregation of the Faith asserts. Certainly it is curious that a document purporting to reflect modern scientific understanding, and with its interdiction of "science without conscience", should repeat the Church's traditional view of the sharp distinction between the human and other species, thereby apparently providing a licence for animal experiments much broader than seemingly societies would now allow. Many loyal members of the Church will now be asking whether the Vatican is leading them up a blind alley from which only humiliating retreat is possible. For outsiders, the important questions are whether the interdictions of the Instruction will impede the fuller understanding of human embryology and whether the Instruction's oblique incitement to civil disobedience in the pursuit of strict legislation carries weight. □

European research crisis

The European Communities' research programme will be decided next week too lightheartedly.

THERE is something in the view that Europe has been living in a crisis of transition since the time of Charlemagne. Europe's problem is that of reconciling its sense that united action will be beneficial with the wayward determination of its several national states that they should individually say what form united action should take. The European Communities have been slaves to this dilemma since their creation 30 years ago. Now the problem afflicts their programme for research. A plan to spend more than 8,000 million European currency units (roughly the same number of US dollars) over the five years beginning in January was put forward last year by the European Commission's chairman, M. Jacques Delors, as a test of Europe's willingness to take the Commission seriously. In the event, the British cancelled the meeting last December at which the issue should have been decided. Next week, in Brussels, the member governments will probably settle for a more modest programme, lopping 1,000 million ECU off the original bill.

Will that mean that the cause of common action has been advanced? Or set back on its heels? All interested parties will no doubt form their own opinions. The sad truth, however, is that these will be out of tune with the opinions of most working researchers in European states that next week's compromise will not presage the evolution of the common research programme that Europe needs. The objection to the Commission's plan, even as modified for next week's decision, is that only some of its components carry the promise of succeeding as the Commission would wish.

Some features of the proposals are splendid. Only the Commission, for example, could have given Europe a chance of contributing significantly to the development of fusion reactors by means of its sponsorship of the Joint European Torus (JET); it is to be hoped that the outcome of that experiment will indeed allow work to begin of the detailed design of NET (Next European Torus) in 1990 as the Commission hopes. The case for this project, which will consume a seventh of the research budget, is that no single member of the Communities would have chosen to mount it from its own resources. But member states have been ready to cede precedence to the Commission because the goal is so distant that national competitive instincts are not directly aroused. But it will not be long before the Commission finds that it has earned a substantial fund of goodwill, and perhaps even a willingness by member states to cede it further responsibilities, for having carried through this project.

The same, unfortunately, cannot be said of other components of the research programme. That for stimulating pre-commercial development in computer technology (called ESPRIT) is big enough, at 2,050 million ECU but not decisively so. Indeed, this spending is only a small part of what European governments and companies are likely to spend in this field over the next five years. Yet it remains to be seen whether the Commission's programme will add something distinctive to the pattern of work elsewhere. More often, the programme appears to those who work with it as a means of augmenting their own sources of finance by means of applications to Brussels. In an ideal world, the Commission would have been required to produce a much more formal and rigorous evaluation of the work already carried out before being given a second tranche of funds.

Much the same applies to its other proposals for the next five years. It is hard to believe that much will come of the Commission's proposals in biotechnology or in the battle against AIDS (although there is a data-collection job to be done by somebody); national governments are too much engaged already. It may for the long run be more significant that the Commission now offers a programme for making the research institutions of its members more flexible, and their people more mobile. □

Europe underestimates cost of space programme

- ESA Paris meeting to review budget
- Trading national space projects

London

THE European Space Agency (ESA) has underestimated by more than £600 million (1,100 million European Currency Units, ECU) the amount needed annually to fund its new space programme, precipitating a major rescheduling of its proposed international research projects.

The British National Space Centre (BNSC) says the cost of the programme will be 2,800 million ECU, a third more than the previously expected cost.

The 13 nations that comprise ESA will need to curtail their space programme substantially, or demand even more money from the members. An ESA council meeting is scheduled for the end of this week in Paris, when members will be asked to declare their commitment to the space plan and to discuss a revamped programme.

ESA had agreed in Rome in 1985 to increase the European space budget by nearly 70 per cent over three years. The basic outline of the programme had also been agreed, at an estimated cost of 1,750 million ECU a year at the peak of the programme in 1990.

There is likely to be considerable friction over the projects to be jettisoned to bring down the cost of the programme. France is keen to develop the Hermes spaceplane and the Ariane heavy launcher; West Germany, Italy and the United Kingdom have given support to the space station project.

BNSC still awaits a response to the 15-year space plan presented to the government last summer (see *Nature* 324, 197; 1986). Attempts were made two weeks ago to get together all the relevant government departments to agree a space budget — expected to be about £200 million a year, with half spent within the United Kingdom — but it failed to take place. Ministerial discussions are still in progress.

British Labour Party science policy

ON Thursday 26 March 1987, *Nature* will be providing a forum in London at which Labour Party representatives will outline their policy on science.

As space is limited, anybody wishing to attend should telephone Neil Crombie in London on 01-836 6633, extension 2480. □

The UK space programme is at present largely funded through the Departments of Trade and Industry and of the Environment and the Ministry of Defence. The BNSC plan would bring all this support under its own control, an option likely to be vigorously opposed by the relevant ministers.

The ESA meeting this week will demand not only that the members make a commitment to Europe's long-term space hopes but will also be one of many to prepare for an ESA ministerial summit scheduled for June.

ESA members have been particularly concerned about developing space for peaceful use, and are still disturbed about US intentions to devote military resources to the space station, previously considered a civil project (see *Nature* 325, 563; 1987).

Norway and Austria became full members of ESA this year, joining West Germany, Ireland, France, Denmark, the United Kingdom, the Netherlands, Spain, Italy, Sweden, Switzerland and Belgium. Finland has associate member status and Canada has a 'special arrangement'.

The shortfall in the ESA budget (costs at 1985 prices) could provoke serious disharmony within the group. Members are keen to pursue space technology, but each project is assessed individually; horse trading will begin this week, with each country trying to ensure its share of the research and development cake.

The ESA members are well aware of the opportunity now afforded them as a result of the mishaps experienced by the US space shuttle programme. The development of the Ariane V heavy launcher, in particular, could offer substantial commercial opportunities to the Europeans for launching civil communications satellites, although fierce competition is expected from Japan, China and the Soviet Union.

Europe has also been keen to promote more use of the European Communications Satellite network. Two satellites have already been launched and carry television and telecommunications signals to most parts of Western Europe.

The race within Europe to gain a foothold in direct broadcasting by satellite is now under way, with France, West Germany and the United Kingdom in the front line. Reliable launchers and space communication technology will be fundamental in ensuring the success of these programmes.

Bill Johnstone

Protropin status questioned by FDA decision

Washington

A DECISION by the US Food and Drug Administration (FDA) may cause the biotechnology company Genentech to lose its position as market leader in sales of recombinant human growth hormone (rhGH). Last week, FDA granted approval to Eli Lilly & Co. to market another version of the hormone in a move that could change the interpretation of the 1983 Orphan Drug Act.

Genentech's rhGH, tradenamed Protropin, is a 192-amino-acid protein, consisting of the 191-amino-acid hormone normally found in humans plus an extra methionine residue at the amino-terminal end, an artefact of recombinant production in *Escherichia coli*.

Protropin was granted approval by FDA in 1985 under the Orphan Drug Act, as a treatment for children with insufficient growth hormone. The act provides financial incentive for manufacturers to develop drugs for rare diseases, where they cannot expect to recoup development costs, let alone make a profit and entitles the manufacturer to exclusive marketing rights for seven years.

But human growth hormone sales could reach \$70–80 million in the next year, and many believe orphan drug status should not have been granted. Last year, all of Genentech's \$43.6 million product income resulted from sales of Protropin. Such success prompted three other companies, Ares-Serono, Bio-Technology General, and Eli Lilly, to apply for marketing approval of variants of rhGH lacking the initiation methionine.

Eli Lilly's drug, Humatrope, is also produced in *E. coli*, but does not bear the amino-terminal methionine. Eli Lilly claims its methionyl-free rhGH is less antigenic than Genentech's, and because it has a different structure, it should also be given orphan drug status. Last week the FDA concurred, and gave Humatrope marketing approval.

Genentech is contesting that approval infringes the seven-year marketing exclusivity it is entitled to for Protropin. On 6 March, Genentech filed suit against FDA in federal court in Washington, to force FDA to clarify how the act applies to genetically-engineered drugs. It also filed a preliminary injunction against Eli Lilly to prevent them from marketing Humatrope, and this hearing is scheduled for 26 March. Genentech is also seeking orphan drug approval for its own methionyl-free version of rhGH. The result of all of this could be a complete redefinition of the Orphan Drug Act.

Carol Ezzell

AIDS becomes a notifiable disease in Japan despite protests

Tokyo

JAPAN is about to introduce legislation to make AIDS (acquired immune deficiency syndrome) a notifiable disease. But critics claim that the law will be a grave infringement of human rights.

The following are the main provisions of the AIDS bill: doctors must report the age, sex and source of infection of AIDS-virus carriers to the prefectural government concerned within 7 days; if a doctor judges that a patient may ignore his instructions, the doctor should immediately report the name and address of the patient to prefectural authorities; the prefectural governors are authorized to recommend or order AIDS tests and medical checkups for people infected or suspected of being infected with AIDS, and a maximum fine of ¥100,000 will be imposed on those who refuse; doctors and public servants who fail to keep the secrets of patients will be subject to maximum imprisonment of up to one year or a fine of less than ¥300,000; foreigners infected with AIDS will not be allowed into Japan, for the time being.

Health and Welfare Ministry officials responsible for drawing up the law have found themselves pulled in two directions. On the one hand, some of their scientific advisers, who are involved in the treatment of haemophilic AIDS patients, wish to guard the privacy of patients; on the other, government officials and politicians are pressing for draconian measures to halt the spread of the disease.

Until very recently, most Japanese regarded AIDS as a disease confined to foreigners and a few Japanese homosexuals and haemophiliacs. But the death

of a female prostitute in Kobe in January and revelations in the press about a female AIDS-carrier in Kochi have sent government officials rushing to draft anti-AIDS legislation.

According to press reports, the woman in Kochi ignored her doctor's advice, became pregnant without informing her husband that she was an AIDS carrier, and insisted on giving birth even though there is a high risk the child has been infected. (The birth was announced on Monday, but blood tests of the infant will take several weeks.)

In an early draft of the law, those infected with AIDS who engage in 'dangerous acts' (giving blood or having sexual relations) in the full knowledge that they are infected would have been subject to a maximum fine of ¥300,000 (\$2,000) or imprisonment for up to one year. And doctors who failed to report AIDS carriers to prefectural authorities would be subject to similar punishment. But protests that these provisions constitute an infringement of human rights led the ministry to drop these clauses from the final draft.

Hoei Ohama, chairman of the recently formed subcommittee on AIDS, of the ruling Liberal Democratic Party (LDP), told reporters that "it is more important to prevent the spread of AIDS than to protect the privacy of high-risk groups" and went on to add that "if we respect the human rights of one person, we are depriving 99 others of the right to live".

Some Justice Ministry officials have even suggested that people infected with AIDS who knowingly pass on the virus to others should face injury or murder charges under existing criminal law.

The Health and Welfare Ministry says the new law will protect the privacy of AIDS sufferers as long as they follow doctors' advice. But what is not made clear is how those secondarily infected with AIDS will be tracked down while preserving that privacy. The clause banning the entry of foreigners is also extremely vaguely worded and there is no description of how it will be enforced.

The National Association of Haemophiliacs has condemned the legislation as "inhuman". But if some members of the LDP have their way the law will be strengthened during Diet debate. Tatsuo Ozawa, leader of an AIDS study delegation to the United States, announced in Washington on Friday that he will recommend compulsory medical checks for those employed in massage parlours, and strengthening of the enforcement of AIDS preventative measures while developing measures to protect human rights.

David Swinbanks

Tests for AIDS in pregnancy on offer to all

Munich

VOLUNTARY testing for AIDS (acquired immune deficiency syndrome) should be available for all pregnant women. That was the conclusion of an international symposium on infectious diseases in gynaecology and obstetrics held at the University of Munich earlier in the month. The bread-and-butter of the conference, which brought together more than 500 gynaecologists from nine countries, were topics such as toxic-shock syndrome and premature delivery.

A pilot programme of AIDS testing among expectant mothers has already begun in the university hospitals of Munich. According to Ernst Weissenbacher, the organizer of the conference and professor of medicine at the university, "almost nobody" has refused to be tested.

Of the estimated 100,000 carriers of the AIDS virus (HIV) in West Germany, between one and seven per cent are thought to be women. Virtually all carriers of the infection are members of high-risk groups. At present, the results of AIDS tests are confidential between patients and their physicians which, according to Weissenbacher, helps to explain why pregnant women are ready to volunteer for tests. But Bavarian physicians may soon be required to report carriers of the AIDS infection to local health authorities in which case, says Weissenbacher, the situation could change.

The government of Bavaria has put forward one of the most far-reaching programmes of AIDS testing yet proposed. Under intended legislation, tests would be mandatory for applicants for public service jobs as well as those "in danger of infecting others", including members of high-risk groups such as prostitutes and intravenous drug users and their associates.

Under the planned law, those found to be carriers of the infection would be required to inform their physicians and their sexual partners of their condition. The Bavarian government's plans are vigorously opposed by the Minister of Health in the federal government, Rita Süßmuth.

Steven Dickman

AIDS education programme in Japan

Tokyo

JAPAN is taking several steps to combat AIDS (acquired immune deficiency syndrome). The Health and Welfare Ministry has distributed a 100-page booklet on the disease to prefectural governments across the country; the Education Ministry has instructed education boards to initiate AIDS education from the primary school level up; and television coverage on the disease is now extensive. AIDS-related stocks have soared on the Tokyo stock exchange.

Now Tatsuo Ozawa, a former Health and Welfare Minister who led a mission to the United States last week to study AIDS, has proposed the formation of a private foundation to produce education films and video tapes and to maintain a free phone-counselling service.

David Swinbanks

New director at BGS

London

THE UK Natural Environment Research Council (NERC) has appointed Mr F.G. Larminie to a two-year term as director of the British Geological Survey from September, on the retirement of the present director Mr G.I. Lumsden.

□

Cattle farming in the desert heat an unexpected success for Arabs

Riyadh

HOLSTEIN cattle, originally bred to suit the temperate climate and green pastures of the Netherlands, would not seem to be the ideal choice for a dairy herd in the scorching Saudi-Arabian desert. But 17,600 Holsteins on a farm 50 miles south of Riyadh, the country's capital, not only survive regular summer temperatures of 45–50°C, but are good milk-producers.

Even more remarkably, what is claimed to be the largest herd of cattle in the world is surrounded by 'greened' desert that produces about two-thirds of the cattle feed. And tucked in among the cattle pens is a processing plant that processes the daily 125,000 litres of milk produced by the herd, turning much of it into laban, the 'liquid yogurt' that is so popular in Saudi. The only snag is that the business is so costly that even its ordinary milk has to sell at about £1 a litre.

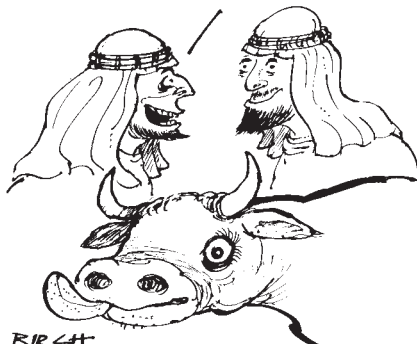
Begun with some assistance from Alfa-Laval, the Al Safi Dairy Farm is one of the numerous businesses of Prince Abdullah Al Faisal, eldest son of the late King Faisal. It is close to the town of al-Kharj, in a region that has been so effectively greened in the past few years that Saudi now exports wheat rather than importing most of it. What has made the transformation, of course, was water and irrigation.

As it happens, there is no real shortage of water. "Drill anywhere in the region", says one Briton employed on the farm, "and there is a 99 per cent chance of finding water." The problem is that most of the water has around 2,500 parts per million of salt, resulting in kidney problems for the cows, a milk yield that is probably considerably less than it would otherwise be, and the need for saline-tolerant crops.

With very little, and no reliable, rainfall, the 5,000 m³ of water needed each day comes from wells. But the original aquifer is drying out, necessitating wells to an aquifer at about 1,800 m. Not only is that water too salty to be potable but it also has to be cooled from 75°C to below 40°C before it can be sprayed on crops.

The water is distributed continuously to each field by a pivot irrigation system,

...and in a few years, we'll have our own butter mountain!



similar to those responsible for the large circular fields seen from the air over parts of the United States. Around fifty 100-acre fields are cultivated in this way, growing mainly alfalfa, which can be cropped every 15 days and is used both as green feed and as a source of hay. Oats and sorghum are two other major crops.

The farm is impressive in its scale and integration and its water supply seems assured. But the cost and quality of the water, and doubts about the ability of the land to continue supporting such intensively grown crops, raise questions about the long-term success of the project.

Peter Newmark

New science minister

Munich

WEST Germany's new minister for education and science, Jürgen Möllemann, is likely to encourage the development of privately run universities that compete with state-supported higher education. Möllemann, 41, has been a paratrooper as well as an elementary school teacher. He is the first education minister from the Free Democratic Party (FDP) and one of four FDP ministers in Prime Minister Helmut Kohl's new cabinet, one more than the last cabinet.

The FDP has long been in favour of allowing market forces to operate in German higher education and has helped to establish privately funded chairs at state-run universities. But almost all the 52 private colleges and universities in West Germany offer specialized curricula not available in the state-run educational system and therefore do not offer direct competition.

Introducing a competitive element will not be easy. The costs of establishing a new university are so high that some state support is essential. The tendency has been for the German Länder (states), which share jurisdiction over higher education with the West German government, to refuse aid on the grounds that sufficient programmes are already offered in the state-run universities.

S.D.

Turk researchers harassed

London

A TURKISH research ship, the *Piri Reis*, belonging to the Marine Sciences Institute of the "Ninth September" University (Izmir) has become involved in the dispute between Greece and Turkey over drilling rights in the Aegean. The *Piri Reis* has made a number of geological survey trips both in the Aegean and the Black Sea. According to a report from the Anatolia Press Agency, Greek naval vessels shadowed the *Piri Reis* when it put to sea on 28 February. On the night of 2–3 March, the agency says, a Greek "attack vessel" with all lights extinguished approached within 30 m of the *Piri Reis* and suddenly switched on all floodlights, to the alarm of the Turks. But the captain of the *Piri Reis*, Metin Kircadagl, and Dr Sukran Dirik, the head of the scientific team, declared that research would continue as long as weather conditions permitted.

V.R.

Frank Press re-elected

Washington

MEMBERS of the National Academy of Sciences have re-elected Frank Press for a second six-year term as president.

Press, 62, first became president in 1981, succeeding the late Philip Handler. He has been a member of the academy since 1968. Earlier, he served as President Carter's science adviser.

J.P.



THIS colour photograph of the recent supernova (SN 1987a) in the Large Magellanic Cloud (LMC) (see p. 239) was taken at the European Southern Observatory in La Silla, Chile on 25 February, two days after the supernova exploded. The super-

nova, which had reached a visual magnitude of 4.5 at the time this photograph was taken, is the lower right (round) of the two bright objects above and to the left of the main body of the LMC. The other, more diffuse object is the Tarantula Nebula. □

Change of heart signalled in Soviet-US health agreements

Washington

FOR the first time since 1978, representatives of the US and Soviet Union biomedical and public health fraternity will meet to exchange ideas. The exchanges had been regular in the mid-1970s but ceased in the wake of the Soviet invasion of Afghanistan. This time, representatives of both countries will meet for five days on the campus of the National Institutes of Health (NIH) in Bethesda to put an official stamp of approval on a new list of projects.

Prospects for collaboration are bright—est in three areas: cancer, heart disease and arthritis. Claud Lenfant, director of the National Heart, Lung and Blood Institute, was a member of the delegation led by NIH director James Wyngaarden last November and has been a regular visitor to the Soviet Union. Peter Fischinger, deputy director of the National Cancer Institute, also went on the visit.

Another area of interest for both countries is Lyme disease. It is named after the town where it was first identified in the United States, and Lawrence Shulman, director of the new National Institute of Arthritis and Musculoskeletal and Skin Diseases, says there is great interest to learn whether it is the same as tick-borne erythema seen in the Soviet Union. Both are associated with ticks, but the Soviet disease seems limited to a skin disorder, whereas Lyme disease is often accompanied by either acute or chronic arthritis, and sometimes both.

The thaw in biomedical relations has been under way for some time. Last fall, two separate high-level US delegations visited the Soviet Union. A month before the Wyngaarden trip, Surgeon General C. Everett Koop headed a delegation that discussed potential exchanges in the public health arena. A speech by President Reagan in 1984 first generated the impetus for the visits, but they were delayed by differences over the treatment of Andrei Sakharov and his wife Yelena Bonner.

Two joint agreements cover all biomedical and health exchanges between the United States and the Soviet Union, one for medical science in general signed in 1972 and the other for artificial heart development signed in 1974. The agreements run for five years and are automatically renewed unless either side objects.

Even during the years of decreased contact, collaborative efforts in arthritis have persisted. Last spring, the *New England Journal of Medicine* published a joint study between researchers in the two countries on penicillamine and hydroxychloroquine as treatments for juvenile rheumatoid arthritis. Earl Brewer of

Baylor College of Medicine in Houston, first author on the arthritis study, says another study using oral gold therapy will be presented at a meeting this summer, and there are new plans of a study using methotrexate as a treatment. Brewer says the arthritis studies have shown that joint protocols can be made to work, and can provide useful information.

Wyngaarden hopes that the example set by the arthritis study can carry over to evaluations of cancer therapies, speeding the development of new treatments. But although some facilities in the Soviet Union are quite modern, others must get along with primitive equipment, says



Windom — new hope for health?

Wyngaarden. He also says communication both to and within the Soviet Union can be difficult. "We seem to know more about some of their institutes than they do", he says.

Some studies are politically easier to attempt in the Soviet Union than others. Studying fat or salt in the Soviet diet has been acceptable, but collaborative efforts to measure alcohol consumption would not be attempted. Political sensitivity is not restricted to the Soviets. Jack Schmidt, acting deputy director of the Fogarty International Center at NIH, says Soviet charges that AIDS (acquired immune deficiency syndrome) is somehow related to US biological weapons development "could have an impact on the extent of cooperation", particularly with regard to AIDS. Soviet interest in AIDS is on the increase (see *Nature* 326, 5; 1987).

Although US officials see the forthcoming meeting as a positive step, they would prefer that formal mechanism for collaborations were not necessary, something the Soviet Union has insisted on.

The joint committee on health will meet on 13–17 April at NIH. Robert Windom, assistant secretary for health and Professor Oleg Shchepin, deputy minister of health, will head the two delegations. The choice of Shchepin came as something of a surprise to US officials who were expecting Yuri Isakov, deputy minister of health for international cooperation, to be chosen.

Joseph Palca

Respite in sight for India's battered frogs

New Delhi

INDIA, the world's largest exporter of frogs' legs, has banned the trade with immediate effect. Conclusive evidence from the Ministry of Environment and Forests and the Indian Council of Medical Research that frogs play a major role in the control of agricultural pests and mosquitoes has precipitated the ban. Frogs have now been placed in Schedule 2 of the Wildlife Protection Act of 1972, giving them a special protected status.

Export of frogs' legs from India, mainly to Europe and the United States, jumped from 390 tonnes in 1950 to 4,065 tonnes in 1982. An estimated 60 million frogs are caught each year from paddy fields and are processed for export. The frog trade provides jobs for 160,000 people and earns India about \$10 million in foreign exchange.

One reason for the export ban is the increase in the pest population that attacks the rice crop. Ninety per cent of frogs' food consists of agricultural pests, including caterpillars and crabs notorious for damaging rice seedlings. It has been estimated that catching frogs for export leads to the survival of 200,000 tonnes of pests and insects, thereby requiring farmers to spend more on pesticides. India imports each year pesticides worth about \$100 million, several times as much as the income from frogs' legs.

Frog catching is done at night, and frog hunters cause serious damage to rice fields. The resurgence of malaria and recent epidemics of Japanese encephalitis, which killed more than 3,000 people, are also partly blamed on the export of frogs, which are thought to restrict mosquito-breeding.

The export of frogs' legs, like that of rhesus monkeys, has been a contentious issue in India for some time. The government has been trying to conserve frogs through such measures as banning export during the breeding season (May to August), restricting the size of frogs for processing and introducing a quota system for export from different regions. It has also asked the countries of the European Economic Community to stop importing frogs' legs from India. But large-scale exploitation has continued.

The total ban on export has been welcomed by those campaigning against conditions in frog-processing centres. The Indian ban is unlikely to alter restaurant menus in Western countries so long as Bangladesh and Indonesia continue to export frogs' legs.

K.S. Jayaraman

NASA launches wind-tunnel tests by Cray-2 supercomputer

San Francisco

A MILITARY brass band played and experimental aircraft circled as NASA (National Aeronautics and Space Administration) officials in California formally unveiled on 9 March the world's most powerful supercomputer complex for designing aircraft.

At a dedication at the NASA-Ames Research Center, 40 miles south of San Francisco, NASA chiefs hailed the Numerical Aerodynamic Simulation (NAS) system as ranking in importance with the first use of a wind tunnel in 1873 by Francis Herbert Wenham and the first flight of the Wright brothers in 1903.

For NASA, the new development is, at the least, a badly needed sign that the post-Challenger space agency has not entirely lost its knack as a pathfinder in advanced engineering. Among the most formidable tasks for the new facility will be the design of the X-30 prototype for the National Aerospace Plane, a hypothetical hypersonic vehicle endorsed by the Reagan administration.

With NAS, aeronautical engineers expect for the first time to be able to derive

The NAS complex is the first routinely to use "Reynolds-averaged Navier-Stokes" equations to chart in three-dimensional detail turbulent supersonic airflow over complex aircraft bodies. Its colourful animated displays can be commanded to appear at a score of workstations at Ames and at dozens of others linked by high-speed data links at other NASA centres, aerospace companies, defence laboratories and universities.

The powerful mathematics of fluid dynamics have previously been limited to highly idealized configurations, such as isolated and geometrically simple airfoils, and often to two dimensions. The Navier-Stokes equations have been called "innocuous in appearance . . . marked by insidious pitfalls", with more than 60 partial derivative terms when expressed in three dimensions. The capacity of 10^9

operations a second expected at NAS by 1989 should make it possible to calculate flow through a grid of a million points arrayed around aerodynamic shapes.

William F. Ballhaus Jr, one of the originators of NAS, now says the facility took two major gambles. One was to order the Cray-2 in 1984 when NASA approved major spending, even though the machine did not then exist. The other was to settle on the AT&T UNIX operating system, now a common protocol in managing computer networks but not when NAS was laid out. As the Ames team guessed right, the powerful machine has been easily incorporated in interactive networks among distant laboratories.

In spite of the new computer power, wind tunnels will continue to hold a vital place in aircraft design. Ballhaus says that "a computer can optimize the design, . . . but to get its actual performance out to the third decimal place, we will still need to fly a real model in real air". And then, somebody must climb into a real airplane and take off in it.

Charles Petit

Soviet Academy of Sciences make room for next generation

London

THE Soviet Academy of Sciences last week adopted a policy of "rejuvenation". In order to make room for younger talent, nobody over the age of 65 will normally be allowed to hold an administrative post in any of the institutes, laboratories or departments belonging to the academy — a ruling incidentally that bars 65-year-old Andrei Sakharov from such a job. But superannuated academy scientists will be permitted, indeed encouraged, to devote their attention to training their successors. "In the interests of preserving intellectual potential", the new post of "honorary director" of an institute will be created, and a "staff of counsellors" made up of senior members of the academy is also envisaged.

The rejuvenation policy was formally presented and approved at the annual general meeting of the academy after several months of discussion. Last October, when Academician Vitalii Ginzburg, head of the I.E. Tamm Department of Theoretical Physics at the Lebedev Physics Institute reached his 70th birthday, he published an article in the monthly *Priroda* advocating a policy of honourable retirement for academy scientists. The Soviet Academy has for years been heavily over-burdened with septuagenarians and octogenarians (in May 1985, according to Ginzburg's figures, more than 55 per cent of full members of the academy were over 70, and 68.7 per cent over the new 65-year-old limit).

The academy, as the annual general

meeting made clear, faces a massive efficiency drive. All fundamental science will be put on a footing of long-term forecasting and planning — a task that, Academy Vice-President Vladimir Kotelnikov noted, will have to be carried out soon and with limited resources. The academy presidium has defined 135 lines of research deemed to be of "immense significance to the national economy". Information science, the synthesis of new materials and scientific instrument-making are to receive priority attention, Kotelnikov said, as "in these spheres difficulties arise over obtaining information from capitalist countries".

Academy departments and research establishments are to become more autonomous, which will free staff of much routine administrative work, but which will doubtless entail some interim confusion.

Scientific research in general must be made more "dynamic" and, Academy President Gurii Marchuk stressed, fundamental research must not simply support but must "overtake" the demands of technology and production. There is to be a major development of science in the east of the country, with the establishment of far eastern and Urals branches, which may well mean some scientists being encouraged to move eastwards.

In such an atmosphere of change and upheaval, many senior members and employees of the academy may well find honourable retirement a welcome option.

Vera Rich



The NAS system's Cray-2 supercomputer

practically relevant solutions of the hydrodynamic equations for the flow of compressible viscous fluids developed more than a century ago and known as the Navier-Stokes equations. The solutions will be used as tools in roughing out designs of aircraft whose performance overwhelm ordinary computers, and which often cannot be tested even in the most powerful wind tunnels.

A Cray-2 supercomputer is, for the time being, the heart of NAS. The cooled machine is at least eight times faster than machines now used in computational aerodynamics and its memory is 40 times bigger. It is not the world's only Cray-2, but it has features giving it more versatility than any other.

In a year or so, NASA officials here expect to install a machine four times as fast. Front-runners for the sale are the still-embryonic Cray-3 or Cray YMP from Cray Research Corporation.

Howard Hughes Institute and tax collectors bury the hatchet

Washington

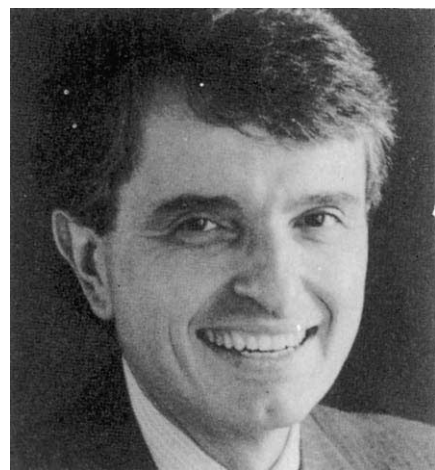
THE Trustees of the Howard Hughes Medical Institute have agreed to pay \$35 million in back taxes without protest as part of the settlement of a long-standing dispute with the Internal Revenue Service (IRS) over Hughes' tax status.

In addition to the \$35 million, to be paid by the end of this fiscal year, Hughes has agreed to spend \$500 million over and above the minimum amount required to preserve its status as a medical research organization (MRO).

Hughes and the IRS have debated for some time exactly what MRO status means. Established in the 1950 tax code, MROs must be involved in "direct conduct of research in conjunction with hospitals", a phrase Hughes president Donald

Fredrickson says has been studied "with Talmudic discipline". Very few foundations in the United States have chosen to use the MRO classification, certainly none with endowments approaching the \$5,000 million of Hughes, so the rules were not clear. Fredrickson says the nature of biomedical research has shifted since 1950. Much important biomedical research is being conducted at institutions not connected with hospitals.

Hughes was anxious for the flexibility to become involved in areas not so tightly tied to clinical research. The settlement with IRS appears to make that possible. Joseph Perpich will join Hughes on 15 April to direct a new programme that will take Hughes activities into new areas of making grants and offering fellowships.



Perpich — expanding Hughes' horizons.

George Thorn, former president of the Hughes Institute and now chairman of the board of trustees, says the major emphasis will be on education. A preliminary meeting to discuss how much money is available and how it will be spent was scheduled for earlier this week in Washington.

Losing MRO status would have meant becoming a private foundation, something the institute was anxious to avoid. Foundations must adhere to more complicated rules in distributing money, and are more restricted in the type of assets they may possess, says Thorn. Private foundations must also spend 5 per cent of their endowment annually, as opposed to the 3.5 per cent required of MROs. Fredrickson says reverting to being a foundation would have required massive organizational changes in the structure of the 24 facilities the institute operates and the 1,000 people it employs. Estimating tax penalties for failing to achieve MRO status had also loomed as a potential nightmare.

Private foundations have, however, been able to survive under IRS rules. Richard Lyman, president of the Rockefeller Foundation, says he regards the 5 per cent spending figure as reasonable. Favourable economic conditions have permitted Rockefeller's endowment to grow from \$1,200 million to \$1,800 million this year.

The Hughes Institute's fortunes changed dramatically with the sale in 1985 of the Hughes Aircraft Company to General Motors. That sale provided the bulk of the institute's endowment, and transformed it into a major force among private philanthropies. In addition to the medical institutes it operates, and the new grant-giving endeavour, the Hughes Institute is also supporting a major effort to establish a database for human genetic information at Yale University. This new database may become the *de facto* standard for information collected by a project to sequence the entire human genome now being considered by several countries.

Joseph Palca

Sizewell B reactor gains expected approval from UK government

London

PREDICTABLY, the British government has given official approval to the construction of a £1,500-million pressurized-water reactor (PWR) at Sizewell on the Suffolk coast. The station, which will join a gas-cooled nuclear reactor on the same site, will be the first of five nuclear complexes, costing £6,000 million in total, installed around the United Kingdom.

The approval met criticism from both opposition parties; they raised doubts about the safety aspects of the PWR in the wake of Chernobyl, and the economic case for the reactor in the light of current fuel prices.

The Energy Secretary, Mr Peter Walker, was undeterred. He concluded that "there is good confidence that Sizewell B is sufficiently safe to be tolerable and that the national need for the station overrides the local interest in favour of conservation".

The reactor has been the subject of much public debate since it was first suggested in the late 1970s. A public inquiry was held between January 1983 and March 1985 from which Sir Frank Layfield, the chairman, was able to produce a report published in January of this year.

The Layfield study recommended approval, although it conceded that comprehensive tests on the safety of the design could not be carried out at this stage. The opponents of the PWR called for better assurances and claimed that the report was irrelevant because it took no account of the accident at Chernobyl.

Walker was adamant that the programme will go ahead. He said: "The

PWR design for Sizewell B is of a different reactor type from the Soviet RBMK design (Chernobyl). All nuclear power stations in the UK, unlike those in the USSR, must have engineered control and automatic protection systems. Moreover, our system of regulation, unlike that



which applied in the USSR, ensures that there is a proper and reliable procedural framework of controls."

The experience of the United Kingdom, he continued, demonstrated that "there is a superior safety culture" to that at Chernobyl. Many of the conservationists who gave evidence against Sizewell B do not share the minister's view and plan a public rally, to be staged in London, to protest at the government's decision. **Bill Johnstone**

Kazakh Academy scientists face party leaders' fury

London

THE former First Secretary of the Communist Party of Kazakhstan, Dinmukhammed Kunaev, whose replacement last December by an ethnic Russian, Gennadii Kolbin, triggered two days of nationalist disturbances in Alma-Ata, faces a party inquiry on charges that include corruption, nepotism and "ethnic favoritism". This is the second step in his public disgrace (he was expelled from the All-Union Politburo in January), and there is already speculation that it may end in a formal trial. Similar allegations have also been levelled against the scientific and scholarly community of the republic. The Minister of Higher Education, Kupzhasa Nari-baev, has resigned after publicly acknowledging his faults, the leadership of the Academy of Sciences of the Kazakh SSR has been severely criticized by party leaders in Moscow, the former head of the science and education department of the Kazakh Central Committee has been severely reprimanded and a shake-up of the universities is under way.

The demonstrations began with students from Alma-Ata University taking to the streets. According to some reports, the banners carried by the demonstrators ranged from "Kazakhstan for the Kazakhs!" and "Kolbin go home!" to demands for a separate representation for Kazakhstan at the United Nations, and even for union of Kazakhstan with China.

The ethnic situation in Kazakhstan is complex. Immigration (often involuntary) from the European territories of what is now the USSR goes back to the nineteenth century, while in the 1950s the "virgin lands" policy of Nikita Khrushchev brought in a new wave of non-Kazakhs. As a result, Russians outnumber Kazakhs by 41 per cent to 36 per cent of the republic's population, with the remaining 23 per cent made up of some 20 other Soviet nationalities. But, under Nari-baev, an unauthorized policy of positive discrimination in favour of ethnic Kazakhs had been in operation. Soon after the disturbances,

Hocus pocus

THERE seems to be in wide circulation a letter purporting to be from *Nature* and referring to a paper entitled "Neoplasms as uncontrolled regenerative hyperplasia" by Cris Garcia Sarmiento (deceased).

We should like to make it clear that this letter is a hoax, that we have no knowledge of either the paper or the author, and that other organizations as well as ourselves have been victims of this hoaxer.

Editor, *Nature*.

the Soviet media noted that the student ringleaders came for the most part from southern Kazakhstan, where there is least ethnic mixing. From this area, too, came ex-minister Nari-baev, and so, it appears, do six of the ten rectors of higher educational establishments in Alma-Ata.

The Academy of Sciences of the Kazakh SSR was also affected by the same failings, according to a report by the party control committee of the Central Committee in Moscow. The committee found that the academy had failed to ensure "the correct representation of scientists of different nationalities" among its members, that there was an atmosphere of "irresponsibility, obsequiousness, mutual flattery and total licence" and that the academy had wasted state funds on "ostentatious mass events", receptions, banquets and gifts to visiting notables.

Even worse, it had failed to carry out its scientific duty as coordinator of research and in the acceleration of scientific progress in the republic. Within the academy's research institutes, the control committee found, the standard of fundamental research had fallen considerably, the amount of original development work had declined, no licences had been sold and 45 per cent of the "solutions" produced turned out to be neither innovative nor useful.

None of the academy institutes had joined the new "scientific and technical complexes" in the Soviet Union to obviate bureaucratic delays in technical innovation. Nepotism and local preference had

Super conductivity

New Delhi

A TEAM of researchers at the Tata Institute of Fundamental Research in Bombay and the Bhabha Atomic Research Centre at nearby Trombay claims to have demonstrated the onset of superconductivity in a synthetic oxide at a temperature of 120 K as well as zero electrical resistance at 90 K. The team says that it has thus surpassed the record of superconductivity at 82 K set by a group led by C.W. Chu at the University of Houston (see *Nature* 325, 756; 1987).

Although the materials used by the two groups are very similar, the Indian team claims that their success in finding a higher transition temperature arises from "sample stoichiometry, heat treatment and preparation technique". The team used samples of yttrium/strontium/barium copper oxide which were annealed in an atmosphere of oxygen and powdered and cast into disks of 8-mm diameter. A month ago, the team found zero electrical resistance in a similar compound with lanthanum instead of yttrium. K.S. Jayaraman

led to the appointment to senior scientific posts of poorly trained and "morally lax" people.

The main responsibility for these shortcomings was attributed by the control committee to the former president of the academy, Askar Kunaev, who was replaced last year. During his term of office, said the report, Kunaev "crudely flouted the principles of collective leadership", ignored critical observations and advice from the All-Union Academy in Moscow, drank to excess and often failed to come to work or to deliver important speeches. Vera Rich

SERC admits funding crisis with no respite in the years ahead

London

THE Science and Engineering Research Council (SERC) has formally admitted that the new salary increases agreed with academic staff last month will have a dramatic effect on its research grant budget. The new salaries will cost an extra £7.8 million in the year beginning 1986 and the following year will rise to £10 million or about 'one-tenth of the annual research grant expenditure'.

SERC is concerned that the full weight of these increases must be borne by its current budget. A similar shortfall of some £20 million because of the increase in the European Organization for Nuclear Research (CERN) due to currency fluctuation, provoked the council to apply this year to the Advisory Board for the Research Councils (ABRC) for extra funds.

The disclosures are outlined in the

council's new three-year plan, *The forward Look, 1988/89 to 1990/91* published last week, which concludes that: "There are important areas of basic and strategic work which Council will not be able adequately to exploit".

The announcement of the SERC shortfall coincides with statements made by the organization called Save British Science which has called government attention to a deepening crisis within the research councils. Scientists are uncertain, they claim, whether their current applications will even be considered, and are fearful that their work will be "arbitrarily terminated".

Meanwhile, a working group commissioned by the government will advise on applications of defence research and development to the commercial sector.

Bill Johnstone

Problems of Portuguese science

SIR—As a Portuguese scientist working for twelve years in West Germany and associated with the academic society of Portugal for five, I read the survey on Science in Iberia (*Nature* 324, 313–332; 1986) with amusement: your journal has been used very effectively for the international airing of proposed programmes of the Portuguese Ministry of Education and its institutions. Whether these programmes will ever become a reality only time will tell.

I do not wish to question the good intentions of those concerned, but their effectiveness. This government has been in office for over a year. The real problems of science in Portugal are: (1) the legal structure of its institutions and their hiring policy (people are rarely fired) and (2) the attitudes and scientific competence of the majority of people in these institutions.

The first problem concerns the ministry and its power to introduce reforms. Why, for example, do teaching assistants with the equivalent of a master's degree effectively have tenure and a guaranteed permanence in the educational system? Why is there so much inbreeding in Portuguese educational institutions? Why are academic positions in Portuguese institutions almost never advertised in the international scientific press, where they would be brought to the attention of interested and qualified Portuguese living abroad? Why is it possible for Portuguese academics to have more than one job?

The second problem has historical roots which have both affected the Portuguese attitude towards science and the world and filled the Portuguese institutions with an astonishing number of incompetent people. It is well known that any multi-step process, such as an educational or scientific establishment, is only as efficient as its least efficient step. There are some excellent and internationally competitive people and research groups in Portugal, but they are relatively few. Progress is, more often than not, hindered by the almost uncountably large number of dog-in-the-manger types who still prevent the opening of Portuguese scientific culture to the rest of the world. This is partly because they feel intimidated by science at a modern international level and partly because they no longer speak the same language. Admittedly, the Ministry of Education is not directly to blame, but some initiative for change should, nevertheless, come from there. Why, for instance, is there no legislation enforcing a periodic evaluation of scientific and teaching competence of

Portuguese scientists and professors as there is in almost every other progressive country? Why is demonstrated excellence never materially rewarded? Why are students in Portuguese universities never asked to evaluate course or entire degree programmes offered to them?

Finally, some specific comments regarding statements in your survey: Every Portuguese is very used to hearing catchy terms like "new blood" (p. 332), which makes it all the more difficult to understand why inbreeding is still a way of life in Portuguese institutions of higher education. "Patience" (p. 330) with all officialdom in Portugal is a necessary requisite for survival in that country. It is not the authorities that require this commodity.

WINCHIL L. C. VAZ

Max-Planck-Institut
für Biophysikalische Chemie,
Abt. Molekulare Biologie,
Postfach 2841,
D-3400 Goettingen, FRG

Citation counts

SIR—The recently reopened debate about citation counts (*Nature* 324, 95; 1986 and 325, 478; 1987) has highlighted the fact that we still do not know enough about citation behaviour to endorse the use of such analysis without qualification in matters of resource allocation. The technique aims at quantifying the impact of published research, this being one measure of the research output of university departments. Output should not, and cannot meaningfully, be considered in isolation from input, however.

A recent study conducted at Lancaster has investigated the determinants of research output in departments of economics in 40 British universities. A bibliometric measure of output per staff member (call it X_1) was constructed for the years 1980–84. This was expressed as a function of university and department size, staff age, staff–student ratio, library stock, external funding and numbers of professors and research students. More than 70 per cent of the variation in X_1 can be explained in this way. The regression residuals — which represent that part of output that cannot be explained by differences in the inputs defined above — could themselves conceivably be used as a performance indicator. Call this latter measure X_2 .

The Spearman's rank correlation coefficient obtained by comparing X_1 and X_2 is -0.597 . This indicates no significant degree of positive correlation between the two measures. If this is the case for economics, one shudders to think at how misleading citation counts might be for sciences where resource costs are greater.

Presumably the aim of performance indicators is to estimate value added. Citation counts alone are quite incapable of doing this. Even if bibliometric analysis could crudely reflect research output, they alone could never measure productivity.

GERAINT JOHNES

Department of Economics,
Gillow House,
University of Lancaster,
Lancaster LA1 4YX, UK

New Zealand denial

SIR—In an article "Genetic manipulation living outside regulation" (*Nature* 324, 202; 1986), it is stated that a team from Oregon State University in Corvallis had field-tested a modified vaccinia virus in sheep, cattle and chickens in New Zealand. The facts are that a scientist from Oregon State University came to New Zealand and requested permission for a field test of the vaccinia/sindbis organism, but this request was denied. After consideration of the request and careful examination of the experimental protocol by the Wallaceville Animal Research Centre Biological Society Committee, in consultation with the New Zealand Advisory Committee on Novel Genetic Techniques, tests were allowed in calves and chickens under strict quarantine. No field trials were permitted and as far as I am aware there have so far been no field trials of any genetically engineered organisms in New Zealand.

J.N. PARLE

Wallaceville Animal Research Centre,
Private Bag,
Upper Hutt, New Zealand

University rows

SIR—I read with interest your opinion on professional athleticism at Oxford (*Nature* 325, 470, 1987).

The boat race between the eights of the Universities of Pavia and Pisa takes place each year alternately on the waters of the Ticino and of the Arno, the rivers flowing through the two old university towns. The rule that the rowing men must be students has in the past caused a hiatus of some years in the annual event when the presence in both crews of oarsmen who would not know what the social sciences were was considered too much, even given the Italian genius for compromise.

We did not at the time indulge in the luxury of brooding on such trivial matters but the survival of the two universities was then as always threatened by lack of funds. The University of Oxford is to be envied for being able to enjoy such luxuries when short of funds.

MARCO FRACCARO

Collegio Fratelli Cairoli,
Università di Pavia,
27100 Pavia, Italy

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

Neutrinos from the supernova

Physicists appear to have reacted as promptly as the better-drilled astronomers to the opportunity presented by the explosion of a supernova in the Large Magellanic Cloud on 23 February.

THE first few weeks of Supernova 1987a have seen a flurry of IAU (International Astronomical Union) telegrams, hastily written preprints and late night telephone calls. Astronomers, already accustomed to dealing with unexpected celestial events, have mobilized with admirable speed. It is even more remarkable that physicists at underground detectors around the world have reacted with equal speed to the reported detection of neutrinos from the explosion. Now, as the supernova itself continues to shine steadily in the Large Magellanic Cloud (LMC), some of the confusion of the first few days is being dispelled.

The detection of neutrinos has created much excitement. The announcement by the Italian experiment in the Mont Blanc tunnel that neutrinos had been seen was soon followed by a conflicting report from Japan. Physicists at the Kamioka zinc mine also saw a burst of neutrinos, but four and a half hours later than the Mont Blanc events. Because a supernova can emit neutrinos only at the very instant of explosion, these reports could not both be right. Yet at the outset it was thought that no other experiment would be able to see the neutrinos (see *Nature* 326, 11; 1987). In particular, the high-threshold IMB (Irvine-Michigan-Brookhaven) device, situated in an Ohio salt mine, can see neutrinos only above 20 MeV, where the thermal spectrum of neutrinos from a supernova explosion should produce almost nothing.

While the discordant observations from Italy and Japan were being pondered, the IMB collaboration announced that it had in fact seen neutrinos from the supernova, and with a clear and unambiguous signal. Starting at UT 7:35:41.37 on 23 February, eight events at an energy between 20 and 40 MeV were recorded over a period of six seconds. To within a few seconds, this is the time the neutrinos arrived in Japan, but as the Kamioka experiment does not have accurate absolute timing, the two can be regarded as simultaneous.

Although this agreement was gratifying, there was a problem. Kamioka saw just one event at 35 MeV, while IMB saw eight events of comparable energy. After a little thought, this turned out not to be a difficulty. A supernova does indeed produce fewer neutrinos at higher energy, but this is partly compensated for by the increase of neutrino interaction cross-section with the square of the energy:

high-energy neutrinos are therefore more likely to be detected and for these particular neutrinos, the IMB detector is effectively seven times more sensitive than that at Kamioka.

In all this excitement, the Italian detection has rather been forgotten. The general opinion is that it saw a chance cluster of events, only a little above background noise. At both IMB and Kamioka, data have been re-examined at the time of the supposed Italian events; nothing has been found.

Meanwhile, as the euphoria of this first achievement of extragalactic neutrino astronomy subsides, physicists are trying to extract more subtle information from their data. Theoretically, at the moment of core collapse, a supernova will emit a burst of neutrinos lasting for about a second. Yet IMB and Kamioka report neutrino events over periods of six and thirteen seconds, respectively.

One explanation is that neutrinos have mass: massive neutrinos travel at less than the speed of light, so that the arrival times of neutrinos of different energy would be spread out. A simple estimate suggested that a mass of some tens of electron volts would disperse the neutrinos over a period of six seconds during their passage from the LMC. But there is a more prosaic explanation. Rather than coming directly to us from the supernova's collapsed core, the neutrinos would most probably have been scattered a few times by atoms in the stellar atmosphere on the way out, their arrival delayed by a few seconds.

In principle, the effects of neutrino mass and simple scattering can be distinguished by careful study of the energy spectrum and arrival times of the neutrinos, because the two effects depend differently on energy. People at IMB and Kamioka hope that their data will shed some light on this question, but it is unclear whether the neutrinos detected worldwide are enough to establish any useful spectral information.

The mere detection of neutrinos is proper cause for jubilation, but astronomers have been in this business longer and have more exacting standards. They continue to build up their picture of the supernova, comparing that with those of previous supernovae observed in galaxies more distant than the LMC. According to Robert Kirshner, of the Harvard-Smithsonian Center for Astrophysics, SN 1987a must be categorized as Type II be-

cause it has hydrogen lines in its emission spectrum, but is not altogether typical.

SN 1987a is fainter than one would expect, its light output rose more steeply than usual and has then quickly levelled out. What will happen next is not quite clear. The light curve of a typical Type II supernova rises to a maximum and then begins to fall in a matter of days. But SN 1987a does somewhat resemble two supernovae observed in the 1940s, which, after a plateau of constant brightness, brightened again before subsiding.

Different Type II supernovae may behave differently because the early and late periods of the light curve are differently controlled. The luminosity of the initial outburst derives from the mechanical energy of the explosion, but in the later stages the energy source is the decay of radioactive nuclei, especially isotopes of cobalt synthesized in the explosion. Unusual ups and downs in the light curve may thus reveal something about the original explosion, the progress of nuclear fusion in the outburst, and the chemical composition of the progenitor star.

This progenitor has not yet been identified. The star Sanduleak-69.202, close to the explosion site, was the first and would have been the ideal candidate — except that observations by the International Ultraviolet Explorer satellite show that it is still there, as indeed are all the known stars in the vicinity. But Kirshner is not greatly concerned. He believes that the progenitor was a slightly fainter star, near the line of sight to 69.202 and obscured by it. Even so, it is irksome that the first nearby supernova for several centuries should have another star right in front of it.

SN 1987a may nevertheless be cosmologically important. It is close enough for Kirshner to hope that speckle interferometry will reveal the angular size of the expanding shell over the coming weeks and months. Knowing the brightness, expansion speed, size and distance, it will be possible accurately to reconstruct the explosion and the subsequent expansion, which in turn will help similarly to reconstruct supernovae in more distant galaxies. Because careful observations of extragalactic supernovae can provide good estimates of their distance, the potential payoff is that SN1987a will help to fix more accurately a few points on the still sparsely occupied plot of galactic redshift against distance.

David Lindley

The strategy for selecting HPRT⁻ stem cells described by Hooper *et al.* in this issue⁴ was somewhat different. This group, at the University of Edinburgh, has developed a method for maintaining ES cells in the undifferentiated state in the absence of feeder fibroblasts using conditioned medium from buffalo rat liver cells⁷. This procedure makes culture and selection of the stem cells much easier, and also opens up the possibility of identifying factors required both for the growth of the stem cells and for the regulation of their differentiation. By adding 6-thioguanine to the conditioned medium of an XY cell line the authors selected spontaneous HPRT⁻ mutants already present in the population. One of these was used to make germline chimaeras, in collaboration with Alan Handyside and colleagues at the MRC Experimental Embryology and Teratology Unit in Carshalton, and with Marilyn Monk at the MRC Mammalian Development Unit in London, who has developed microscale methods for measuring HPRT activity in single embryos and hair follicles. The pattern of inheritance of the *Hprt*⁻ gene from the founder chimaeras is essentially the same as that reported by Kuehn *et al.*. Crossing an HPRT⁻/HPRT⁻ carrier female

with a normal male generates both HPRT⁺ male, and normal and heterozygous female embryos, all of which can be distinguished by enzyme assay at the morula stage. This augers well for early prenatal diagnosis of HPRT⁺ XY human embryos by measuring enzyme activity in single blastomeres sampled from 8- or 16-cell morulae.

The male HPRT⁺ mice generated by both groups are anxiously being watched for signs of behavioural abnormality, but so far they appear perfectly normal. Current ideas about the origin of the neurological symptoms in Lesch-Nyhan humans centre around the observation that HPRT activity is about five times higher in certain regions of the brain (the basal ganglia) than in other cells. Absence of HPRT activity in the basal ganglia appears to be associated with poor development of the terminal arborizations of dopaminergic neurons. It may be just too early to see symptoms in the HPRT⁺ mice, because in humans the most dramatic effects do not develop until several months after birth. Alternatively, there may be subtle differences in purine or neurotransmitter inter-

mediary metabolism between mice and humans which means that mouse neurons escape damage. Pinpointing these differences could lead to the design of new drugs. Meanwhile, clever schemes are being devised for selecting other mutations in X-linked and autosomal genes expressed in ES cells, and for rapid screening of pools of infected cells for retroviral insertion into genes for which selection is not so simple. The full potential of the technique will be realized with the development of methods for replacing normal endogenous genes with specific mutations by homologous recombination. □

1. Lesch, M. & Nyhan, W.L. *Am. J. Med.* **36**, 561–570 (1964).
2. Stout, J.T. & Caskey, C.T. *A. Rev. Genet.* **19**, 127–148 (1985).
3. Milewski, E.A. in *Recombinant DNA* **9**, 88–130 (US Dept Health Hum. Serv. Tech. Bull., 1986).
4. Hooper, M.L., Hardy, K. Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292–285 (1987).
5. Kuehn, M.R., Bradley, A., Robertson, E.J. & Evans, M.J. *Nature* **326**, 295–298 (1987).
6. Robertson, E., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445–448 (1986).
7. Smith, A. & Hooper, M.L. *Dev. Biol.* (in the press).

Brigid Hogan is head of the Laboratory of Molecular Embryology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

Natural selection

Competition among different taxa

Jared M. Diamond

As species of the same genus have usually, though by no means invariably, some similarity in habits and constitution, and always in structure, the struggle will generally be more severe between species of the same genus, when they come into competition with each other, than between species of distant genera^[1].

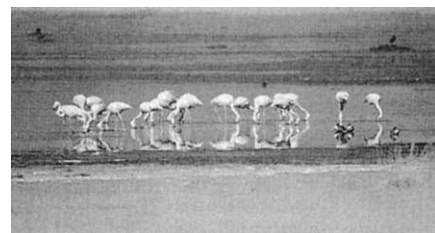
FOLLOWING this opinion of Darwin, ecologists have tended to focus on competition between closely related species, especially members of the same genus. But a critical resource can also be shared by taxonomically remote species. Antarctic krill, for example, sustain baleen whales, seals, penguins, petrels, fish and squid, not to mention krill-harvesting ships. Thus, competition between divergent taxa could be significant. Recent studies pose questions about how taxonomic closeness affects the consequences of competition.

Inferences of competition can be based either on indirect evidence (such as shared limiting resources and complementary spatial and temporal distributions) or on direct evidence (such as observations of aggression and effects of removal experiments). These types of evidence are illustrated by Carpenter's field studies² of Californian flowers whose nectar is sometimes in short supply and on which hummingbirds, bees, moths, butterflies and flies depend. Hummingbirds chase bees and moths from flowers, whereas bees in turn chase hummingbirds. When humming-

birds are present, visits to flowers by moths decline because hummingbirds chase some moths away and interrupt the foraging bouts of the rest. Moth visits peak at noon in the absence of the birds, and at dawn or dusk in their presence (hummingbirds forage during the day). When Carpenter enclosed flowers in cages that excluded hummingbirds but let bees in, visits to the flowers by bees became more frequent and lasted longer.

Birds and bees are also the subject of a recent study³ by Schluter on the effects of competition between bees and Darwin's finches visiting flowers in the Galapagos Islands. Finches spend 20 per cent of their time collecting nectar on islands lacking bees, but only 4 per cent of their time on islands with bees. In the absence of bees the finches shift in body size towards the smaller bees, from weights of 14–21 g on bee islands to 10–12 g on islands without bees. On the islands without bees, individual finches that visit flowers are smaller (average weight 8.7 g) than individuals not visiting flowers (11.3 g). The explanation is that nectar feeding is more economical for bees and small birds than for large birds, because they thereby satisfy a higher fraction of their energy needs.

Hurlbert *et al.* have recently obtained evidence⁴ that flamingoes compete with fish for zooplankton at high-altitude lakes in the Peruvian Andes. The number



Flamingoes at Laguna Parinacochas, Peru, feed on the abundant zooplankton in the absence of fish. (Courtesy of Stuart Hurlbert.)

of flamingoes per lake varies from 0 to 4,457. Fish predation on zooplankton reduces zooplankton biomass ninefold in lakes with fish compared with lakes without fish, and shifts the zooplankton composition from large-bodied to small-bodied species. As a result, flamingo numbers are 10 times lower at lakes with fish than at lakes without fish. Most of the instances of lakes entirely without flamingoes turn out to be lakes with fish.

In the light of such studies, how might competition between distant taxa differ from that between members of a genus? Competing members of a genus often exhibit mutual aggression; mutual displacement of each other in space and time; and mutual character displacement (evolutionary changes) in morphology in each other's absence. The more dissimilar the competitors, the more asymmetrical the three phenomena are likely to be. Hummingbirds, for example, chase moths and not vice versa, and the sheer numbers of the smaller moths may swamp the birds by forcing them to devote so much time to chasing that foraging becomes uneconomical ('aggressive neglect'; see ref. 5). Trapping seed-eating rodents causes a much greater increase in abundance of seed-eating ants than the increase in abundance of rodents caused by poisoning the ants⁶.

The generality of these predictions requires further testing. They gain interest because of the possible economic importance of competition between humans and other animal species and the implications for conservation. Thus, hunters blame wolves for depleting deer and fishermen blame sea birds for depleting fish. More likely, however, the asymmetry of competition is in our favour: human starvation caused by effects of sea birds has yet to be documented, but massive starvation of sea birds has certainly resulted from overfishing by humans⁷. □

1. Darwin, C. *On the Origin of Species* (Murray, London, 1859).
2. Carpenter, F.L. *Am. Zool.* **19**, 1105–1114 (1979).
3. Schluter, D. *Am. Nat.* **127**, 95–102 (1986).
4. Hurlbert, S. H. *et al. Limnol. Oceanogr.* **31**, 457–468 (1986).
5. Ripley, S. D. *et al. Am. Nat.* **93**, 127–134 (1959).
6. Brown, J. H. in *Community Ecology* (eds Diamond, J. & Case, T. J.) 41–61 (Harper & Row, New York, 1986).
7. Mills, S. *New Scientist* **91**, 10–13 (1981).

Jared M. Diamond is Professor of Physiology at the University of California School of Medicine, Los Angeles, California 90024, USA.

Intracellular signalling

Receptor-operated Ca channels

Erwin Neher

THERE is hardly any tissue where receptor-operated calcium channels have not been postulated. Evidence from biochemical and tracer flux studies indicates that they are involved in many diverse cellular functions (see ref. 1 for a review). In contrast, electrophysiologists have had until recently only one well-characterized mechanism to explain receptor-mediated changes in Ca conductance: the second-messenger-induced modulation of voltage-activated Ca channels, which has been found in several systems²⁻⁴. No such voltage-activated channels, however, have been found in cell types such as mast cells⁵ that are believed to possess receptor-operated Ca channels. But in the past few months, three more mechanisms of receptor-mediated Ca flow have been suggested. One is described by Kuno, Gardner and colleagues (ref. 6 and page 301 of this issue⁷); another was reported⁸ by Reuter and colleagues; and a third was proposed⁹ by Irvine and Moor.

Inositol (1,4,5) trisphosphate (InsP₃) is well established as a second messenger that releases calcium from intracellular stores^{10,11}. The report in this issue⁷ describes a second, independent effect of InsP₃ on calcium influx. In T lymphocytes, the authors observe current fluctuations resulting from opening and closing of very small, Ca-permeable channels directly activated by intracellular InsP₃. In their earlier work⁶, they observed a very similar channel in human T lymphocytes after activation by mitogen. The channel is not or is only marginally voltage-dependent and is blocked rather than activated by increasing calcium concentration, [Ca]_i, on its cytoplasmic side. The opening of this channel therefore cannot be caused by increased [Ca]_i, which shows that this is not an indirect consequence of Ca release from intracellular stores.

Such channels, which mediate what might be called a Ca-induced Ca influx, have recently been described⁸ by Reuter and colleagues. These authors observed a rise in [Ca]_i following peptide-induced neutrophil activation and the consequent opening of two types of nonspecific channels that are permeable to Ca and other cations. These channels are opened only by increased [Ca]_i and not by InsP₃ or by the peptide directly. This mechanism of Ca-mediated activation of Ca-conducting channels could be a widespread phenomenon that is used by cells to perpetuate an InsP₃-induced increase in [Ca]_i, which by itself is only transient.

The third mechanism of receptor-

mediated calcium inflow through the plasma membrane which has recently emerged also involves InsP₃. Irvine and Moor have shown⁹ that inositol (1,3,4,5) tetrakisphosphate (InsP₄), derived from InsP₃ by phosphorylation, may serve as a messenger on its own (see the recent discussion in News and Views¹²). When injected into sea urchin eggs, InsP₄ causes a raising of the fertilization envelope, a process which is dependent on the presence of extracellular Ca. This result suggests an InsP₄-activated channel.

The three mechanisms described (Ca-permeable channels directly opened by InsP₃; channels opened by an initial InsP₃-induced increase in [Ca]_i; and putative channels opened by InsP₄) will probably help to elucidate the role of calcium in many diverse cellular responses. It is now becoming clear that a phase of sustained increase in [Ca]_i dependent on external Ca can be separated from a transient phase independent of external Ca (mediated by InsP₃) in many types of electrically inexcitable cells. The channels described here have properties suggesting that they mediate such a sustained phase. Although they probably are the receptor-operated Ca channels that have long been searched for,

they are not Ca channels in the traditional sense. They lack ionic specificity, so that under physiological conditions the ion contributing most to the current is not Ca.

At first glance one is reluctant to accept three new mechanisms. But only a limited body of data is available at the moment. It will be interesting to see how many separate mechanisms will remain once they are more completely characterized. It is conceivable that the effect of InsP₃ on lymphocytes, which now seems to be a direct action of InsP₃, is in fact caused by InsP₄ generated by the cell from InsP₃. This possibility is considered by Kuno and Gardner⁷. On the other hand, Irvine and Moor mention⁹ that the action of InsP₄ is enhanced by prior Ca mobilization. This may provide a link to the channels described⁸ by Reuter and colleagues. □

1. Bolton, T. B. *Physiol. Rev.* **59**, 606-718 (1979).
2. Reuter, H. *Nature* **301**, 569 (1983).
3. Kameyama, M., Hescheler, J., Hofmann, F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **407**, 123-128 (1986).
4. Dunlap, K. & Fischbach, G. J. *Physiol. Lond.* **317**, 519-535 (1981).
5. Lindau, F. & Fernandez, J. M. *Nature* **319**, 150-153 (1986).
6. Kuno, M. & Gardner, Ph. *Nature* **326**, 301-304 (1987).
7. Kuno, M., Goronzy, J., Weyand, C. M. & Gardner, Ph. *Nature* **323**, 269-273 (1986).
8. von Tscharner, V., Prod'homme, B., Baggiolini, M. & Reuter, H. *Nature* **324**, 369-372 (1986).
9. Irvine, F. & Moor, R. M. *Biochem. J.* **240**, 917-920 (1986).
10. Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67-69 (1983).
11. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
12. Michell, R. H. *Nature* **324**, 613 (1986).

Erwin Neher is in the Department of Membrane Biophysics, Max Planck Institute for Biophysical Chemistry, Postfach 2841, 3400 Göttingen-Nikolausberg, FRG.

Archaeology

New perspectives on the history of ancient Maya civilization

Jeremy A. Sabloff

OUR understanding of ancient Maya civilization has undergone huge changes in recent years. Although the exciting breakthroughs in deciphering Maya hieroglyphic writing and in appreciating the ideological roles of the Maya elite have justifiably received much publicity¹, equally significant breakthroughs have been made in understanding the agricultural and adaptive foundation upon which Maya civilization was built^{2,3}. All these advances have caused significant changes in traditionally held views of the rise, development, and so-called 'collapse' of Classic Maya civilization. Thus, in contrast to older views, Classic civilization is now thought to have been composed of urban centres, supported by both extensive and intensive agriculture, closely linked economically, politically and ideologically, and subject to widespread raiding and conflict. Whereas Mayan civilization was considered unique, it is now believed

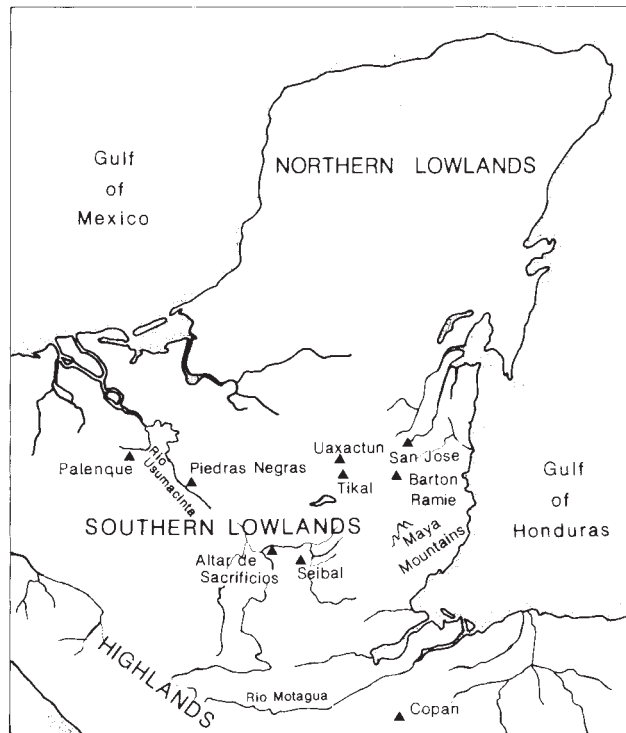
to have developed in a similar manner to other pre-industrial complex societies. The new pollen data presented by David Rue on page 285 of this issue⁴ should lead to further revisions of the traditional picture. Assuming Rue's interpretation of the pollen in his two samples is correct — and it appears both plausible and well argued — some re-examination of current views is called for.

For many years, the earliest Maya remains of settled farmers could not be securely dated before about 800 BC. Although many suspected that earlier remains existed, there were no data to support such suspicions until the major discoveries⁵ by Norman Hammond at the site of Cuello in Belize, which pushed this date back about 1,500 years to 2300 BC. Hammond has argued that the occupants of Cuello in the late third millennium BC were settled, pottery-making, maize agriculturalists who culturally were Maya. But

it remains an open question whether the occupants of Cuello migrated there from outside the Maya lowlands, bringing along with them their agricultural and ceramic skills, or whether they had a long lowland history as well as one of agriculture and pottery-making. Richard MacNeish's recent work in Belize led him to argue⁶ for a long, pre-ceramic, lowland occupation.

Rue's new pollen-core data from western Honduras described in this issue⁴ broaden our understanding of this question. The data provide evidence for slash-and-burn agriculture, probably part of a hunter-gatherer strategy. Given the current lack of ceramics dated to pre-800 BC in western Honduras, Rue's results suggest that there were pre-ceramic peoples practising agriculture there as far back as the fourth millennium BC. This finding indicates that agriculture existed long before the rise of complex societies in the lowlands and did not have a much shorter history of development, as previously believed. To confirm this result, it is necessary to find archaeological remains that can be directly tied to Rue's pollen data. Such evidence will elucidate the nature of this early agricultural occupation, its relation to sedentary and/or mobile adaptations and its possible links with later fully sedentary, village agriculturalists. It is to be hoped that fieldwork directed to this end will soon begin both around Lake Yojoa, near Palenque and elsewhere in the lowlands (see map).

Moving from the foundation of Maya civilization towards its later stages, the 'collapse' of the Classic Maya has also been the focus of much attention. Recent



research indicates that earlier notions of cultural collapse and demographic abandonment in the southern lowlands are not tenable, and that some sites and regions were not abandoned or were only partially depopulated. Furthermore, there were significant demographic movements and shifts of economic and political centres during the ninth century AD (refs 7,8). Rue's pollen discoveries make a useful contribution to the consideration of the demographic and cultural changes at this crucial time by indicating first, that there was very widespread clearing of forest cover in the region around the great Classic Maya city of Copan (see map) at the time of the collapse of this centre; and second, that there was occupation in this

region for several centuries after the abandonment of Copan.

First, the pollen evidence for major forest clearance when Copan stopped being an urban centre will support the position⁹ that ecological stress and soil degradation were important in the downfall of many centres in the southern lowlands during the ninth century AD. Although such degradation may not have been present around all large sites and was probably not the sole cause for the abandonment of many centres, ecological catastrophe at certain key sites could have led to disruptions among centres with close economic and political ties. This could have made cities in the northern lowlands, such as Uxmal, Sayil and Chichen Itza, which began to flower at this time, much more attractive to the Classic Maya population in the south and to non-Classic Maya traders on the western frontiers of the Maya lowlands.

Second, the continuing occupation of rural lands outside Copan fits in well with various new data indicating that the demographic situation in the southern lowlands by AD 900 and after was much more complex and varied than previously thought. Some large centres, particularly in northern Belize, suffered no disruption at all; some showed major population loss; some, like Copan, were abandoned but had a continuing rural population; and others were totally abandoned.

As we begin to obtain a clearer picture of the demographic pattern throughout the lowlands during and immediately after the ninth century AD, it will soon be possible to examine the relationships between this pattern and other factors such as soil degradation and access to water, raw materials and crops like cotton and cacao. Then, a new, testable explanation of the great changes in the Maya world between AD 800–1000 should be possible. □

Graham Goddard (1938–1987)

PROFESSOR Graham Goddard, who made important contributions to the study of synaptic plasticity in the mammalian brain, died in a hiking accident in New Zealand on 15 January 1987.

Born in England in 1938, he spent most of his career in Canada before moving in 1981 to the University of Otago as Chairman of the Department of Psychology. After graduate work at McGill, he taught at the University of Waterloo and it was there, while investigating the role of the amygdala in avoidance learning in rats, that he made the discovery which established his reputation: in several regions of the brain low-intensity stimulation repeated daily eventually leads to the development of epileptic seizures. Kindling, as Goddard called this apparently permanent increase in excitability, continues to be widely studied as a stimulus-

induced model of epilepsy.

More recently, first at Dalhousie University and then in New Zealand, Goddard turned his attention to long-term potentiation, a form of synaptic plasticity found in the hippocampus which provides a plausible model for information storage in the brain. His work in this field led to a series of discoveries of the first importance, including the demonstration of associativity, inhibitory suppression and heterosynaptic depression.

Graham Goddard brought exceptional analytical powers to the study of brain mechanisms, together with a psychologist's awareness of the importance of locating these mechanisms in the context of behaviour. His untimely death has deprived neuroscientists of an original and distinctive voice.

P.O. Andersen & T.V.P. Bliss

1. Schele, L. & Miller, M.E. *The Blood of Kings* (Kimbell Art Museum, Fort Worth, 1986).
2. Turner, B.L. II & Harrison, P.D. (eds) *Pulltrouser Swamp: Ancient Maya Habitat, Agriculture, and Settlement in Northern Belize* (University of Texas Press, 1983).
3. Adams, R.E.W. *et al.* *Science* **213**, 1457–1463 (1981).
4. Rue, D.J. *Nature* **326**, 285–286 (1987).
5. Hammond, N. *et al.* *Nature* **260**, 579–581 (1976).
6. MacNeish, R.S. *Second Annual Report on the Belize Archaic Archaeological Reconnaissance* (Peabody Foundation for Archaeology, Andover, 1981).
7. Chase, A.F. & Rice, P.M. (eds) *The Lowland Maya Postclassic* (University of Texas Press, 1985).
8. Sabloff, J.A. & Andrews, E.W.V. (eds) *Late Lowland Maya Civilization: Classic to Postclassic* (University of New Mexico Press, 1986).
9. Rice, D.S. & Rice, P.M. *Latin American Research Review* **19**, 7–34 (1984).

Jeremy A. Sabloff is University Professor of Anthropology and the History and Philosophy of Science, University of Pittsburgh, Pennsylvania 15260, USA.

Chemical oceanography

Global pollution measured by lead in mid-ocean sediments

Clair Patterson

ON page 278 of this issue¹, Veron and colleagues at CNRS and at Bordeaux University report their studies of the difficult, complex matter of lead pollution in the uppermost layers of outer deep-water sediments in the North Atlantic Ocean. It is common knowledge that heavy-metal pollution prevails within near-shore ocean sediments, and it might therefore be thought that either positive or conclusive negative evidence for lead pollution in the open oceans should also be readily found in pelagic sediments. But this is not the case: the processes that yield the pollution effects easily observable in coastal sediments also limit these effects from extending into the open oceans. Furthermore, mass-sedimentation rates in deep waters far from shore are so slight that there are serious difficulties in studying the very thin, uppermost, decades-old sediment layers involved, although pollution effects are enhanced by these small sedimentation rates. Such problems made investigators in the past reluctant to search for evidence of lead pollution effects in deep-ocean sediments.

Two hundred years ago, lead in the oceans did come from rivers, carried outwards by water transport, but the concentration in deep open ocean waters was then so small that it corresponds to those that now exist in purified laboratory waters (about 10^{-13} g Pb per g water). Today, much more lead contamination of the oceans comes from the atmosphere than from rivers. In earlier times the bulk of natural lead entering the oceans was weathered from continents and was included in silicate and humic particles that were deposited in relatively shallow shelf sediments. The small, aqueous, chelated remainder of river lead was rapidly sorbed

on surfaces of phytoplankton which were then eaten by zooplankton in biologically productive off-shore shelf waters. This biologically processed lead was then deposited in local sediments with zooplankton fecal pellets and phytoplankton debris. The small fraction of soluble lead that escaped these efficient removal mechanisms was advected outwards by water currents to the open oceans, where it was removed from the upper, biologically active, thin, mixed zone by plankton, and finally sedimented in fecal pellets. Some of this lead was released from disintegrating pellets before they settled to the sediment layer, and formed a lower-water reservoir of soluble lead which was recycled upwards to the mixed layer, together with nutrients, to be removed and sedimented downwards again.

Most chemical oceanographers believe that undisturbed natural pre-industrial shapes of depth-concentration profiles of lead in the deep open oceans can no longer be used for direct measurement because of contamination by industrial lead. It is uncertain, for example, whether lead concentrations in the upper 1-km thermocline zones of natural profiles are larger, nearly the same or smaller than concentrations in the underlying 4–5-km-deep water zones.

The timing and magnitudes of changes in lead concentrations in the Earth's atmosphere during recent centuries have been recorded in polar ice strata and were discussed² in a News and Views article last year (Fig. 1). These data indicate that direct inputs of lead carried by rain to open oceans increased enormously in the Northern Hemisphere during the past 50 years. These greater flows of lead from direct atmospheric input routes bypass coastal-sediment trap barriers, which allow only small flows of lead to the open oceans. Today, all measured lead concentration profiles show varying sizes of positive bulges in their upper 1-km-thick thermocline zones (Fig. 2). The bulges of greatest magnitude are located in the North Atlantic, where atmospheric inputs

of lead to the oceans from rain are largest; intermediate bulges are found in the North Pacific, where atmospheric inputs are intermediate; and the smallest bulges are in the South Pacific, where atmospheric inputs are small. This geographical variation in pollution effects results from hemispheric differences among inputs of industrial lead to meridional tropospheric circulation cells. Most industrial atmospheric lead is emitted to the northern westerly winds from North America, Europe and Japan, influenced by a latitudinal west-to-east decline of atmospheric inputs of lead to the oceans caused by short atmospheric residence times of lead-rich anthropogenic aerosols circling the Earth.

Temporal changes in lead concentra-

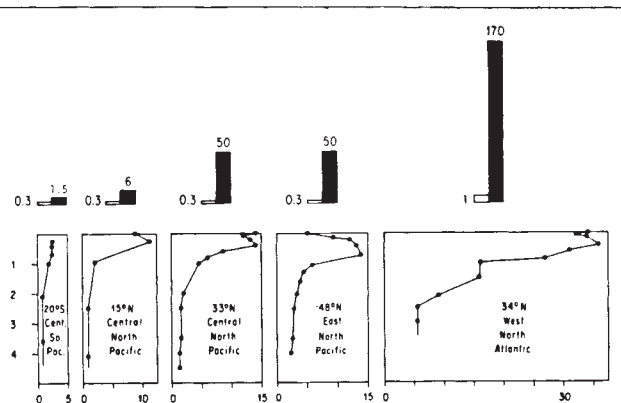


Fig. 2 Comparison of eolian (wind-borne) lead inputs with concentration profiles in the open oceans. Histograms show the net atmospheric lead input fluxes (10^{-9} g cm^{-2} yr), corrected for recycled sea spray. Filled columns, present day; white columns, ancient. Graphs show the lead concentrations of sea water (abscissa; 10^{-9} g kg^{-1}) per km depth (ordinate) (from refs 3, 6, 7).

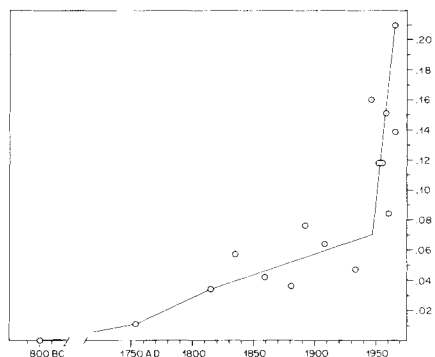


Fig. 1 Increase of industrial lead pollution in Greenland snow with time since 800 BC. Ordinate, concentration of lead (parts per 10^9) per kg snow; abscissa, age of samples (from refs 2, 5).

tions in oceans can be recorded in coral growth layers. Shen and Boyle have recently made the first measurements³ of these changes (dashed curve in Fig. 3) which define the increase of lead/calcium ratios in mid-North Atlantic coral layers during the past century. The congruence of this curve with that in Fig. 1 proves the existence of a close relationship between changes of lead concentrations in the troposphere and in ocean surface-layer waters, and it validates the existing interpretative model for ocean lead depth-concentration profiles. Industrial atmospheric lead emissions increased sharply after 1940 because of the large increase in production and burning of lead alkyls in automobile fuel. Emissions from North America were the source of most of the lead in mid-North Atlantic air before 1970, but the decline in the coral lead curve after that date is probably associated with the decline in alkyl lead emissions resulting from US government restrictions on sales.

Shen and Boyle also measured⁴ stable lead isotope tracers in coral layers, and these data further substantiate the validity of the atmospheric pollution model by showing a close temporal coupling be-

tween changes in the isotope compositions of atmospheric industrial leads and oceanic surface water leads (solid curve in Fig. 3). Before 1880 seawater lead was natural, and possessed a high $^{206}\text{Pb}/^{207}\text{Pb}$ ratio characteristic of fluvial leads weathered from ancient shield drainages of North America. After this year, industrial leads, characterized by low $^{206}\text{Pb}/^{207}\text{Pb}$ ratios in leads of old ore bodies, began to be introduced to the Atlantic from smelter fumes in increasing amounts which, during the next 50 years, gradually reduced $^{206}\text{Pb}/^{207}\text{Pb}$ ratios in successive mixtures of natural and industrial lead in the North Atlantic mixed-surface zone (the residence time of lead in that thin layer is about a year), shown by the decline in the solid curve in Fig. 3 between 1880 and 1940. After 1960 this ratio in North American atmospheric industrial lead increased as a result of a higher proportion of anomalously radiogenic lead mined in Missouri, which was mixed with lead mined elsewhere from much older ore bodies, and used in the production of lead alkyls in the United States. This is reflected by a rise in the solid curve of Fig. 3 between 1960 and 1970. The decline in this curve after 1970 is probably associated with the increased importance of less radiogenic industrial lead emissions from Europe (added to the North Atlantic from easterly winds which then eddy northwards into the westerlies) as the proportion of lead emissions from the United States declined in the mixture of atmospheric leads over the North Atlantic.

It was in this context that Veron *et al.* carried out the measurements of industrial lead in open ocean sediments that they

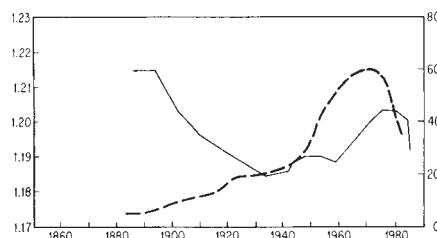


Fig. 3 Temporal comparison of eolian (wind-borne) lead inputs with lead concentrations in corals. Left-hand ordinate, $^{206}\text{Pb}/^{207}\text{Pb}$ ratios (solid line); right-hand ordinate, Pb/Ca in nmol mol^{-1} (dashed line); abscissa, year (from ref. 4).

report in this issue¹. Their data show an excess of lead in the uppermost 400 years of sediment layers in a pelagic core and the uppermost 200 years of layers in a deep outer shelf core (mixing by organisms during the past century extended to these depths) which cannot be accounted for by natural sedimentation processes, whose lead deposition fluxes were determined in contiguous deeper sediment layers. The magnitude of this excess lead equals about half the total integrated input flux of atmospheric industrial lead to the North

Atlantic ocean surface, calculated from input fluxes measured at specific dates normalized to the slope of the Greenland ice lead curve before 1940, and normalized to the lead alkyl consumption curve for the United States, Canada and Europe after 1940. The magnitude of excess deep-ocean sediment lead determined by Veron *et al.* is also approximately equal to the total amount of industrial lead in the North Atlantic thermocline 'bulge', suggesting that about half the industrial lead input is sedimented and the other half resides in the water column. Simplifications involved in these interpretations may have introduced errors but, overall, the find-

ings reported by Veron *et al.* are consistent with, and develop and extend, the prevailing model that explains the concentration of lead in the oceans. □

1. Veron, A., Lambert, C.E., Isley, A., Linet, P. & Grousset, F. *Nature* **326**, 278 (1987).
2. Peel, D. A. *Nature* **323**, 200 (1986).
3. Boyle, E. *et al.* *J. geophys. Res.* **91**, 8573 (1987).
4. Shen, G.T. & Boyle, E.A. *Earth. planet. Sci. Lett.* (in the press).
5. Murozumi, M. *et al.* *Geochim. cosmochim. Acta* **33**, 1247 (1969).
6. Flegel, A. & Patterson, C.C. *Earth. planet. Sci. Lett.* **64**, 19 (1983).
7. Flegel, A.R. *et al.* *Nature* **321**, 689 (1986).

Clair Patterson is in the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA.

Conservation biology

Problems in leaving the ark

Anna Marie Lyles and Robert M. May

UPSTAIRS in the Brooklyn Museum is a beautiful example of a favourite Victorian plaything: a winding line of wooden animals wait in pairs, queueing up to enter Noah's ark. The ark is increasingly being used as a metaphor for conservation programmes, usually based in zoos or aquaria, that aim to preserve species whose continued existence in the wild is threatened by the rising tide of humanity. Contemporary zoos face great problems in embarking their animals, with a foreseeable need to keep 2,000 or more species of terrestrial vertebrates (not to mention other endangered animals) afloat indefinitely¹. This is beyond the capacity of existing institutions, and already poses difficult problems of choice that will get worse as time goes by².

The hope, of course, is that someday, in a better-ordered and less human-centred world, it will be possible to return animals to flourish in something close to their original habitats. In this sense, the concept of the zoo as an ark represents a change from its traditional role as essentially a 'cabinet of curiosities', a museum of living things, with traffic being one way: in and not out. For a few species, someday is now, with deliberate efforts to reintroduce captive-bred animals into their indigenous habitats. In some cases, such reintroductions have followed extinction of wild populations, whereas in others reintroductions have been made to enhance genetic variability or to bolster a declining population. These experiences have led to a growing appreciation of the genetic, ecological and behavioural problems that species may encounter on leaving the ark (ref. 3; M.R. Stanley-Price, personal communication).

One lesson is that the conditions originally causing the wild population to decline should be identified, and if possible remedied, before attempts at reintroduction are made. The story of the endemic

Hawaiian goose, *Branta sandvicensis*, illustrates this point⁴. The birds live in rather inaccessible places, and the causes of their decline to near extinction in the middle of this century have not been adequately identified, much less corrected. About 1,761 captive-reared birds were released between 1960 and 1978, but in 1978 the wild population of about 750 birds — although no longer on the brink of extinction — still was not self-sustaining.

A second lesson is that it is important to choose an appropriate release site. Smaller animals with fairly precise habitat requirements or complex life-cycles may require more exhaustive feasibility studies than do larger generalists. The choice of an appropriate release site can become even more difficult when the species is extinct in the wild. Assessment of its ecological requirements will depend upon inferences drawn from closely related

Eric Hoskins

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hawaiian goose *Branta sandvicensis*.

species, research on captive animals and observations made during pilot-release programmes. The initially unsuccessful programme⁵ to reintroduce the cheer pheasant, *Catreas wallichi*, into Pakistan (where they were thought to be extinct) exemplifies the importance of follow-up programmes. Surplus eggs from British breeders were shipped to Pakistan, hatched and released. Subsequent survival was very low. Radiotelemetry was therefore used to determine the habitat requirements of the birds, which led to the release site being changed and consequently to higher survival.

A third lesson involves a cluster of concepts related to the idea of reintroducing groups of animals whose overall genetic makeup will enhance their success. There seems a good *a priori* reason to suppose that stocks containing intact natural genetic variation will do better⁶. Unfortunately, this variation within the stocks will seldom be available, because captive propagation and reintroduction plans are all too often last-ditch efforts to stave off extinction, and founders are collected after substantial genetic variation has already been lost. Some currently publicized examples, such as the California condor or the black-footed ferret, bring this home.

Given that we rarely have the luxury of reintroducing the full genetic variation found in the original indigenous stocks, there are several ways of selecting compromises that may approximate the correct locally adapted genotype. One idea is to maximize genetic heterogeneity in the released animals by hybrid matings among populations or subspecies. The hope here is that, after release, natural selection will produce the correct 'locally adapted' genotype. This approach has, for example, been adopted for the reintroduction of swift foxes, *Vulpes vulpes*, to the Canadian prairie and for collared lizards in the Missouri Ozarks of the United States^{7,8}. But this hybridization approach can result in disaster. An example is the hybrid ibex, *Capra ibex*, which was reintroduced into Czechoslovakia, and which proceeded to breed in the wrong season⁹.

A different idea is to try to obtain stock for reintroduction from the subspecies or population most likely to have local adaptations, even if this means relatively low genetic variability. Such animals could derive from indigenous stocks (taken into captivity, bred up and reintroduced), or from near neighbours, often a population in a similar habitat. Because animals from different populations are often mixed up in captive collections, this approach is not always possible, but has, for example, been used for reintroduction of the fish *Pociliopsis occidentalis* in Arizona¹⁰.

Even if a captive population starts with a sufficient number of founders to keep intact most of the natural gene pool, and

even if these are kept in naturalistic enclosures where conspicuously unusual selective forces are avoided, loss of the wild genetic variation (that is, domestication) is inevitable if the animals are kept long enough¹¹. It is thus desirable to keep the number of generations between capture from the wild and reintroduction as low as possible. These ideas are illustrated by efforts to preserve the Lord Howe Island wood rail, *Tricholimnas sylvestri*¹². This species declined under the depredations of introduced feral animals, with fewer than 20 birds remaining in 1980. Five pairs were captured and a breeding programme was established on the island. At the end

Eric and David Hoskins

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Looking to a bright future? The otter Lutra lutra. of 1983, 74 birds were reintroduced; the population was back to 200 or more birds in 1984.

A fourth lesson is that behavioural patterns are an important consideration in reintroduction programmes. Social animals should be released as cohesive groups whenever possible, to minimize stress and to facilitate reproduction. Gradual acclimatization of individuals to each other and to the new environment seems the best way to present the groups from scattering on release. This is often achieved by shipping the animals to the release site and keeping them in pens where social groups are assembled. The animals can be fed and provided with shelter for some time after the gates are opened, both to ease the transition for the animals while they learn to use wild resources and also to make monitoring easier.

Such techniques smoothed the reintroduction of the Arabian oryx, *Oryx leucoryx*¹³. Direct persecution by hunting exterminated this animal from the wild in 1972. Fortunately, captive herds had been established during the 1960s and reintroductions into the Oman desert began in 1980, using behavioural data from a closely related species. Seventeen animals were shipped to the Oman and penned as two herds. When social patterns seemed stable and the full range of sexual behaviours was seen, the release was begun. This reintroduction seems to be a success, with 15 live births and only 6 deaths between 1980 and 1984. The project has also gained local support by employing local bedouins as wildlife rangers and drivers.

Reintroductions of territorial animals may be more difficult, as each group may

have to be released at a different site. In England, the otter *Lutra lutra* has recently been reintroduced¹⁴ into patches of its native range from which it was lost in the 1950s and 1960s. Pollution by organochlorine insecticides, coupled with habitat loss and disturbance, contributed to the sudden decline and fragmentation of populations of these otters. Four breeding units (2 females and 1 male in each) were assembled from young animals bred in captivity, acclimatized on-site and released at different sites between 1983 and 1985. These groups seem to have expanded their home ranges successfully and have begun to breed.

The behavioural similarity of released and wild otters suggests that much otter behaviour is innate. This suggests a fifth lesson, namely that animals with flexible behaviour patterns may be very difficult to reintroduce. By analogy to the process of (relatively slow) loss of genetic variability by domestication, culturally determined repertoires of behaviour may be irretrievably lost in zoos. Captive-born animals may be 'rehabilitated' through elaborate training, but there are limits to the skills that can be taught; captive-bred individuals may have limitations to their learning abilities that significantly reduce their viability when they are released.

Victorian Noah's arks and present-day efforts to preserve animal species both focus mainly on mammals and birds, as has this article. We hope to see more concern for endangered species of arthropods and other non-chordate phyla. Insofar as most of these creatures tend to have behavioural patterns that are more 'hard-wired' than those of higher vertebrates, to have higher intrinsic capacities for population growth, and to have relatively small physical sizes (making it easier to keep relatively large populations), their chances of leaving the ark successfully may be higher than for many vertebrates. On the other side of this coin are the very specialized habitat requirements of many of these arthropod and other species. □

1. Soule, M., Gilpin, M., Conway, W. & Foose, T. *Zoo Biol.* **5**, 101-113 (1986).
2. Conway, W.G. *Int. Zoo Yb.* **24/25** 210-219 (1986).
3. Ralls, K. & Ballou, J. (eds) *Zoo Biol.* (special issue) **5**, (1986).
4. Kear, J. & Berger, A.J. *The Hawaiian Goose: An Experiment in Conservation* (Buteo Books, Vermillion, 1986).
5. Ridley, M.W. *Int. Zoo Yb.* **24-25**, 40-44 (1986).
6. Hedrick, P.W., Brussard, P.F., Allendorf, F.W., Beardmore, J.A. & Orzack, S. *Zoo Biol.* **5**, 91-99 (1986).
7. Herrero, S., Schroeder, C. & Scott-Brown, S. *Biol. Conserv.* **36**, 159-167 (1986).
8. Templeton, A.R. in *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soule, M.) 105-116 (Sinauer, Sunderland, 1986).
9. Greig, J.C.S. *Afr. J. Wildl. Res.* **9**, 57 (1979).
10. Vrijenhoek, R.C., Douglas, M.E. & Meffe, G.K. *Science* **229**, 400-402 (1985).
11. Frankham, R. *et al. Zoo Biol.* **5**, 127-138 (1986).
12. Dixon, A.M. *Int. Zoo Yb.* **24/25**, 179-189 (1986).
13. Stanley-Price, M.R. *Int. Zoo Yb.* **24/25**, 179-189 (1986).
14. Jeffries, D.J., Wayne, P., Jessop, R.M. & Mitchell-Jones, A.J. *Mammal Rev.* **16**, 65-79 (1986).

Anna Marie Lyles and Robert M. May are in the Biology Department, Princeton University, Princeton, New Jersey 08544, USA.

Protein structure

New light on old defects

Roger H. Pain

MYOGLOBIN, the 'hydrogen atom of biology', shares with other globular proteins the characteristic that its 1,261 non-hydrogen atoms are remarkably well packed (although not as closely as a box of tennis balls). As I have discussed previously in a News and Views article¹, this relaxed packing is associated with compulsive motion of the atoms in a manner which, like the chemical and physical nature of the protein interior, is anisotropic and which is believed to contribute to the biological function of these molecules.

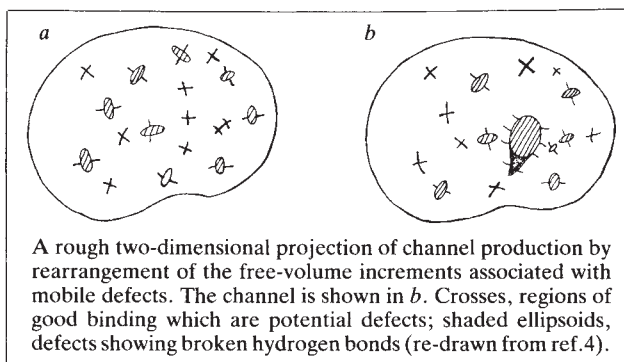
Two quite different experimental approaches, solution compressibility and X-ray crystallography, have now been used by Gekko and Hasegawa² and by Frauenfelder *et al.*³ to shed more light on these basic attributes of protein structure.

There are three kinds of free space within a protein structure: the inevitable vacuum left if and when the atoms are packed ideally to the limit of their van der Waals' radii; spaces of greater than atomic size frequently occupied by water molecules and termed cavities; and defects where atoms and groups are on average separated by subatomic distances. Such defects, or regions of poor bonding interconnected by protein fabric, can provide a high cooperative mobility of free volume and its associated potential energy. These have been called 'mobile defects'⁴ and can provide pathways for the penetration of small molecules to the interior of the folded protein (see figure). The question arises as to whether the defects are the result of the forces that lead to condensation of the polypeptide chain being opposed by thermal agitation, or whether they represent the optimal packing of each particular chain.

Gekko and Hasegawa², using measurements of sound velocity, extend existing data on compressibility to 25 proteins, including myoglobin. All the globular proteins so far studied (with the exception of subtilisin) show a positive adiabatic compressibility β_s , suggesting that the cavities and/or defects can be reduced in size, although quantitation is subject to the uncertainty of the negative contribution of hydration to β_s . This finding is strengthened by the demonstration that values of β_s correlate reasonably well with the independently measured volume occupied in solution by unit mass of protein, V_0 . In other words, the larger the volume occupied by a given length of folded polypeptide chain, the more the

molecule can be compressed.

Because bulk properties must have their origin in the detailed molecular structure it was hoped that it might be possible to identify important features by correlating compressibility with various features of the 25 different proteins. Proteins with a high content of helical structure seem to be more compressible than those with low helix content, although Gekko and Hasegawa knew that correlations could reflect other properties — in this case multi-subunit and multi-domain



structure. Their result could, however, mean that the rather densely packed helical structures are prevented, perhaps by their own rigidity, from packing together in a way that minimizes the defect space. The acid test should be a careful study of the X-ray structures of helical and non-helical proteins.

The correlation of compressibility with the paraffin-like nature or hydrophobicity of the component amino-acid residues reveals a tantalizing result that is not easy to translate into detailed structural terms. Overall hydrophobicity of the protein correlates positively with compressibility whereas, when the content of individual residues is examined, correlation of β_s is observed with only two out of the six major hydrophobic residues. Clearly, some other, unknown property of these residues, possibly associated with their 'packability', is involved.

It is possible to draw conclusions about the dynamic fluctuations of a protein using statistical thermodynamics as originally discussed by Cooper⁵. Gekko and Hasegawa show that the extent of the volume fluctuation is on average 0.3 per cent of the total volume in many proteins. As they point out, this is more than enough, if the increased volume is localized at a few points within the folded protein, to accommodate extraneous molecules in the manner proposed in the mobile defect hypothesis.

Frauenfelder and co-workers³ approach

the question of the basis of defect and cavity space in a different way by examining in detail the atomic coordinates of myoglobin obtained from X-ray crystallographic analysis at 300 and at 80 K. They observe a general shrinkage of the protein crystal for this reduction in temperature, expressed as a 5 per cent reduction in volume of the crystallographic unit cell and a 3 per cent decrease in that of the protein molecule. The fact that cooling was achieved without use of a cryosolvent suggests that the shrinkage is associated primarily with the decrease in thermal energy. Taken together with the compressibility studies, this suggests that a proportion at least of the free space in a protein structure is the result of thermal 'expansion' energy and not just limitation of folding.

Perhaps the most interesting aspect of the work of Frauenfelder *et al.* is the finding that the shrinkage has little effect on the dimensions of the helical cylinders and, similarly, that the proportional shrinkage of the larger cavities is small. In consequence the largest volume change must be occurring in the small packing defects. It is intriguing to postulate, therefore, that it is these defects that are most susceptible to deformation by thermal energy, and by extension

that they, more than other regions, are subject to thermal fluctuations in the protein interior. This could lead on a longer timescale to mobility of the associated potential energy to create transitory larger defects in specific locations into which an extraneous molecule such as oxygen, water or acrylamide could fit¹. Subsequent redistribution of the energy would allow the molecule to move through the protein as easily as a fish-monger through a debutantes' garden party, as proposed¹ in the mobile defect hypothesis.

Boyles' law has been redefined as "the greater the external pressure the greater the volume of hot air". The fact that the formidable amount of work involved in the new studies^{2,3} relating temperature, pressure and volume has not led to instant understanding of protein dynamics shows how unready the protein molecule is to reveal its secrets. But the results do promise fresh insights from the combination of thermodynamic and structural experimentation. □

1. Pain, R.H. *Nature* **305**, 581–582 (1983).

2. Gekko, K. & Hasegawa, Y. *Biochemistry* **25**, 6563–6561 (1986).

3. Frauenfelder, H. *et al. Biochemistry* **26**, 254–261 (1987).

4. Lumry, R. & Rosenberg, A. *Colloques Internationaux du CNRS* No. 246, 53–61 (1975).

5. Cooper, A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2740–2741 (1976).

Roger H. Pain is Professor of Physical Biochemistry at the University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Mathematics

Hilbert's sixteenth problem

Ian Stewart

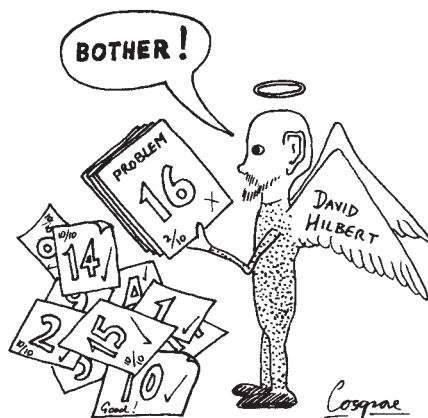
IN 1900 David Hilbert delivered a celebrated lecture before the International Congress of Mathematicians in Paris. "The close of a great epoch," he said, "not only invites us to look back into the past but also directs our thoughts to the unknown future." Hilbert emphasized the importance of specific problems in mathematical research, and listed 23 problems which seemed to him to be of fundamental interest. It is a measure of the maturity of today's mathematics that most of Hilbert's problems now have answers, but some remain stubbornly intractable, the most notorious being the Riemann hypothesis. The sixteenth of Hilbert's problems comes in two similar but largely unrelated parts: the first asks how the different branches of a plane curve can be arranged; and the second raises the analogous problem of the number of limit cycles arising from a differential equation in the plane. There are some important recent developments, not all positive, on this second part of Hilbert's sixteenth problem.

Before describing the current state of play, I will look at the origins of the problem in more detail. In the late nineteenth century, mathematicians were starting to take a serious interest in nonlinear differential equations. In particular they worked on differential equations of the form $dy/dx = Y/X$, where X and Y are polynomials in x and y . In general it is not possible to write down explicit solutions to such equations, but Henri Poincaré discovered that a great deal of information can be obtained by topological means. Rewrite the equation as a system $dx/dt = X$, $dy/dt = Y$, where t is a new variable, thought of as representing time. As t varies, the solution $(x(t), y(t))$ describes a curve in the plane. This curve depends on the initial conditions $(x(0), y(0))$, and by considering all possible initial conditions one obtains a system of curves called the phase portrait of the equation. The phase portrait gives a geometrical picture of all possible solutions to the differential equation, and is a powerful tool for a qualitative study.

In a fundamental re-direction of research on differential equations, Poincaré determined the basic topological properties of such phase portraits. In particular he recognized the importance of limit cycles, which occur when a curve in the phase portrait closes up on itself forming a loop. A limit cycle corresponds to a periodic solution of the differential equation, in which the system repeats the same behaviour over and over again, and hence has genuine physical interest. (In practice, some extra technical requirements are

imposed when defining limit cycles, but I will omit these here.)

The second part of Hilbert's sixteenth problem asks how many such limit cycles can occur, for polynomials X and Y of given degree. When Hilbert raised the question, it was not even known whether the number of limit cycles must be finite. This weaker version of Hilbert's problem appeared to be solved affirmatively when H. Dulac wrote a lengthy memoir (*Bulletin de la Société Mathématique de France* 51, 45–188; 1923). The crucial step in Dulac's proof is to show that an infinite



number of limit cycles cannot accumulate around a single point of the phase portrait.

Some time later, Yu.S. Il'yashenko of Moscow State University (*Uspekhi Mat. Nauk.* 37, 127; 1982) noticed an error in one of the lemmas used by Dulac to establish his main theorem. Now, Il'yashenko has managed to repair some of the results in Dulac's memoir and to prove the finiteness of the number of limit cycles under more stringent conditions. In particular he shows that for polynomials X , Y of degree two, the number of limit cycles occurring in any finite portion of the plane is finite; and that for "almost all" polynomials of degree two, the number of limit cycles in the whole plane is finite. The proofs use modern developments in singularity theory and the theory of normal forms.

A proof of finiteness of the number of limit cycles for polynomials of degree higher than two would be a major step towards solving Hilbert's sixteenth problem. But Hilbert asks for more: namely, the exact number. To be precise, let $H(n)$ be the maximum number of limit cycles that can occur when X and Y have degree n . The known results, described above, cannot even show that $H(n)$ is finite. However, in all known examples $H(n)$ is not just finite, but fairly small. The gap between what is conjectured that exists

and what is known is enormous.

N.N. Bautin (*Mat. Sbornik* 30, 181–196; 1952) found a pair of polynomials X and Y of degree two, for which three limit cycles occur. Thus $H(2)$ is at least 3. This result was bettered by Shi Sonling (*Scient. Sinica* 23, 153–158; 1980) who found an example with four limit cycles. Thus $H(2)$ is at least 4. This example is a small (and very complicated) perturbation of the system given by $X = -y - 10x^2 + 5xy + y^2$, $Y = x + x^2 - 25xy$.

The best results to date are that $H(n)$ is at least $\frac{1}{2}(n^2 + 5n - 14)$ for even $n \geq 4$, at least $\frac{1}{2}(n^2 + 5n - 26)$ for odd $n \geq 9$, and that $H(3) \geq 5$, $H(5) \geq 14$, $H(7) \geq 27$. The first two results here are due to N.F. Otrokov (*Mat. Sbornik* 34, 127–144; 1954). That for $H(3)$ is due to K.S. Sibirskii (*Differents. Uravneniya* 1, 53–66; 1965) and Shi Sonling (*Acta Math. Sinica* 4, 300–304; 1975); the last two follow from a general result of Il'yashenko (*Mat. Sbornik* 78, 360–373; 1969).

Nonlinear differential equations are important in all sciences and are at the forefront of some of the most active areas of current mathematical research. Today's methods for analysing nonlinear differential equations give an excellent understanding of many general phenomena. Moreover, numerical methods, implemented on a computer, can provide very detailed information on any specific differential equation of interest. Hilbert's sixteenth problem falls neatly in between. It concerns a class of equations (those defined by polynomials X and Y of degree n) which is too restricted to be amenable to general methods, yet too broad (having too many free parameters) to yield to numerical analysis. Additionally, Hilbert's problem asks for global information on the entire phase portrait, whereas many of the most powerful general results are only local.

Mathematicians tend to think of differential equations in two variables as being trivial. This is perhaps true of specific equations, or of the local behaviour of arbitrary equations. The status of Hilbert's sixteenth problem illustrates that the global behaviour of multi-parameter families of equations continues to pose almost insuperable problems. We can sympathize with Hilbert's opening words of 1900: "Who of us would not be glad to lift the veil behind which the future lies hidden; to cast a glance at the next advances of our science and at the secrets of its development during future centuries? What particular goals will there be toward which the leading mathematical spirits of coming generations will strive? What new methods and new facts in the wide and rich field of mathematical thought will the new centuries disclose?" □

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Immunization against AIDS in humans

SIR—The discovery of a diversity of subtypes of human immunodeficiency virus (HIV)¹⁻³ immediately raised the possibility of major difficulties in the development of a vaccine against AIDS (acquired immune deficiency syndrome) because a humoral response induced by one HIV strain may generate neutralizing antibodies only against that strain (that is, a type-specific response) and not against other HIV strains. In fact, this is precisely the case in most reported animal experiments⁴. This difficulty prompted us to investigate whether a cell-mediated immune response, based on a signal recognition system that activates both cytotoxic T cells and T-cell-dependent antibody-secreting B-cell clones, might overcome the limitation of type-specific responses to potential immunogens. We wish to report the results of primary immunization of one of us (D.Z.) by one of two vaccines that we believe will stimulate a cell-mediated immune response against HIV infection.

We have previously shown that, after immune activation, cells infected with HIV release virus⁵. Before doing so, they proceed through an immunogenic stage where viral antigens are present on the cell membrane (unpublished results). Activation of a cell-mediated immune response should therefore hinder viral dissemination both by destroying these cells (by means of specific killer T cells) and by the (antibody-mediated) neutralization of free virus. To induce such a response in HIV seronegative immunologically normal individuals, we have used a recombinant HIV-vaccinia virus which has the advantages of triggering a strong cell-mediated immune response and not producing recurrent infection. The recombinant virus^{6,7} contains the *env* gene of a clone of the HTLV-III_B strain⁸ of HIV. The envelope sequence encodes the gp160 protein which includes both the outer gp120 and the transmembrane gp41 peptides. Cells infected with the recombinant virus express gp160 on their membranes⁷.

Preliminary experiments in several species of animals by others^{7,9} and in *Cercopithecus*, baboons and one chimpanzee by us in collaboration with B. Moss, M. Robert-Guroff and R.C. Gallo¹⁰ have demonstrated the safety of the recombinant virus. In addition, our work on *Cercopithecus* has shown the absence of measurable immune impairment during and after immunization.

Primary immunization of D.Z., who is HIV seronegative, involved scarification with 10⁷ plaque forming units of the recombinant virus. Following immunization no systematic reactions (temperature, chills, headache or stiff neck) or regional reactions (enlargement of draining lymph

Table 1 HTLV-III-induced proliferative response and IL-2 receptor expression following immunization

Stimulator	Dilution	Proliferative index		% IL-2 receptor positive cells	
		D.Z.	NNII	D.Z.	NNII
HTLV-III _B	1/50	7	<2	17	<3
	1/100	20	<2	—	<3
	1/200	26	<2	14	<3
	1/400	12	<2	12	<3
	no virus	1	1	<3	<3
	1/20	3	<2	8	<3
UV-irradiated HTLV-III _B	1/50	8	<2	10	<3
	1/200	5.3	<2	5	<3
	1/400	5.8	<2	—	<3
	no virus	1	1	<3	<3

Fresh peripheral blood lymphocytes from D.Z. (63 days after immunization) and from a normal non-immunized individual (NNII) were isolated by Ficoll-Hypaque and suspended in RPMI supplemented with 10% heat-inactivated normal human serum. To aliquots of 10⁵ cells in a final volume of 0.1 ml of medium, were added 0.1 ml of medium containing dilutions of UV-inactivated (2,500 joules) HTLV-III_B or HTLV-III_{RF} derived from a stock solution of live virus (1 µg ml⁻¹ centrifuged from supernatants of infected H9 cells provided by R.C. Gallo). Cells were cultured at 37 °C in a 5% CO₂ incubator. To measure proliferative index, six days after stimulation the contents of each well were labelled with 1 µCi ³H-thymidine for 8 h, after which cell ³H was measured. The mean results of triplicate samples divided by the control value (no virus) are shown. A monoclonal antibody against the Tac antigen (provided by T. Waldman) and a goat antimurine antibody coupled to fluoresceine were used to determine the percentage of IL-2 receptor positive cells. The mean results of duplicates are shown.

nodes or lymphangitis) were observed. A 1.5 cm pustule appeared at the scarification site by day seven and followed the usual evolution of a good response to vaccinia. The immune system was monitored weekly over a nine-week period by measuring CD4 and CD8 cell numbers as well as the *in vitro* production of interleukin-2 (IL-2) from the cells. No immune defect was detected by either assay. The immune response to vaccinia was measured by enzyme-linked immunosorbent assay (ELISA) and showed a 1:2430 titre 30 days after immunization. Antibodies against HTLV-III_B antigens were also detected. On day 30 D.Z.'s serum was highly positive in a competitive inhibitor test for antibodies against gp41. On days 30 and 63 we detected neutralizing antibody¹² against HTLV-III_B with a titre of 1:40 but not against HTLV-III_{RF}, a subtype also referred to as HAT, which diverges from the B strain by over 1,000 nucleotides¹³.

Cellular immune response was studied 63 days after immunization. Fresh peripheral blood lymphocytes stimulated either by UV-irradiated HTLV-III_B or HTLV-III_{RF} were used as responders. Stimulation with HTLV-III_B produced a high level of response as measured by the proliferative index as well as by the number of IL-2 receptor positive cells (see table). HTLV-III_{RF} stimulation also activated the cells but to a lesser extent. A similar degree of activation of the responder cells was observed when fixed HIV-infected autologous cells were used as stimulators.

These results show that immunization of humans with a recombinant vaccinia

virus expressing HTLV-III_B envelope glycoproteins can trigger a primary immune response not only against the immunizing strain but also, albeit to a lesser degree, against a very different strain. Preliminary confirmation of this result has come from immunization of a small group of Zairians, all of whom were HIV- seronegative volunteers and immunologically normal as assessed both by CD4 and CD8 positive cells and by IL-2 production. Two of these individuals developed a significant titre of neutralizing antibody against the HTLV-III_{RF} strain.

To obtain an amplification of the primary immune response, we have recently boosted D.Z. and a few other seronegative volunteers who had received the recombinant virus vaccine. If this boost is not sufficient to produce a high immune response associated with neutralizing antibodies against different HIV strains, alternative approaches — such as the use of multiple recombinant viruses, each expressing a different subtype of HIV gp160 or the use of alternative boosters — will be explored. We would wish to overcome any limitations imposed by subtype diversity before we could properly consider any large scale clinical trial of the vaccine.

While the recombinant HIV-vaccinia virus is our initial choice for immunizing seronegative individuals, for immune deficient AIDS patients we are attempting to induce a cell-mediated response by administration of non-infectious, fixed autologous cells expressing HIV antigens on their surface. We shall report full details of the protocol and the outcome of the

first immunizations of that kind, which have been carried out in Zaïre, in due course. All our work in Zaïre has the full support of the Zaïrian Ethics Committee.

DANIEL ZAGURY
RÉGINE LÉONARD
MICHÈLE FOUCARD
BRIGITTE RÉVEIL

Université Pierre et Marie Curie,
4 Place Jussieu, 75005 Paris, France

JACKY BERNARD
DANIELE ITTELÉ
ALBERT CATTAN

Institut Jean Godinot,
50 Rue Cognac-Jay, Reims, France

LURHUMA ZIRIMWABAGABO
MBAYO KALUMBU
WANE JUSTIN

Cliniques Universitaires de Kinshasa,
Kinshasa, Zaïre

JEAN-JACQUES SALAUN
BERNARD GOUSSARD

Institut National de Recherche
Biomédicale, Kinshasa, Zaïre

1. Wong Staal, F. *et al. Science* **229**, 759-762 (1985).
2. Hahn, B.H. *et al. Science* **232**, 1542-1553 (1986).
3. Starich, B.R. *et al. Cell* **45**, 637-48 (1986).
4. Putney, S.D. *et al. Science* **234**, 1392-1396 (1986).
5. Zagury, D. *et al. Science* **231**, 850-853 (1986).
6. Chakrabarti, S. *Mol. cell. Biol.* **5**, 3403-3409 (1985).
7. Chakrabarti, S. *et al. Nature* **320**, 535-537 (1986).
8. Hahn, B.H. *et al. Nature* **312**, 166-169 (1984).
9. Zarleng *et al. Nature* **323**, 344-346 (1986).
10. Bernard, J. *et al. AIDS Research and Human Retrovirus* (ed. Bolognesi, D.) (Mary-Ann Liebert Incorp, in the press).
11. Robert-Guroff, M. *et al. Nature* **316**, 72-74 (1986).
12. Popovic, M. *et al. Science* **224**, 497-501 (1984).
13. Hahn, B.H. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 4813-4817 (1985).

Cell fusion in viral diseases

SIR—Lifson *et al.*¹ and Sodroski *et al.*² report that syncytia production by the human immunodeficiency virus (HIV) in susceptible cells is a characteristic feature of infection. They suggest that this type of cell fusion is associated with a differentiation antigen (CD4) and requires only the expression of the virus-specific glycoprotein of HIV. The fusion could contribute directly to the development of the specific cytopathology.

During a study of syncytia produced by the parainfluenza 3 viruses in malignant KB cells, we have observed that during the first replicative cycle up to 30 nuclei can fuse in a syncytium at the end of the eclipse period and before the release of mature virions^{3,4}. Polykaryocytes are only generated at low multiplicity of infection; microcinematographic analysis indicates that they result from the extension of the fusion process from infected to normal cells⁴.

In a more recent series of experiments using temperature-sensitive mutants (ts) of vesicular stomatitis virus (VSV) in rat transformed XC cells, we observed that pronounced polykaryocytosis is obtained at the nonpermissive temperature as long as the selected mutant is able to express the envelope-associated G antigen⁵.

Such extensive polykaryocytosis can be generated from a single infected cell⁶ because the glycoprotein propagates through intercellular communications from cell membrane to cell membrane in the absence of any viral replication. Occasionally, the primary infected cell fuses to the proximal side of the membrane of its closest neighbour, leaving the distal side temporarily free⁶. Our results have been confirmed by Florkiewicz and Rose⁷, who demonstrated the fusing capacity of cloned G antigen integrated in L-cells.

When mice are inoculated intracerebrally with the ts VSV mutants, a spongiform encephalopathy can occur⁸. The lesions are disseminated in the grey matter, both in the brain and the spinal cord. The mutant virus multiplies at a very low level in the brain and not at all in the spinal cord, and can be detected neither by infectivity assay nor by electron microscopy. The labelled dendritic processes are swollen, resulting in a characteristic vacuolization marked by peroxidase-labelled antibodies to the G antigen. They also show areas of abnormal configuration of the plasma membrane, with alternating dark and light lamina⁸.

Polar ice test of the scale dependence of G

SIR—A broad class of quantum gravity theories predicts the existence of additional, finite-range, Yukawa-like interaction terms in the gravitational potential at the newtonian limit¹. Because the departures from newtonian behaviour produced by these terms are small and thought to be significant at ranges of a few metres to a few kilometres, any deviation from the classical form of gravity will be difficult to detect in the laboratory, so that experiments performed on a geophysical scale are a useful test of modern unified theories. One way in which a departure from classical behaviour can be visualized is as an apparent dependence of the newtonian gravitational constant, G , on the distance separating the attracting masses. A difference of nearly one per cent between the value of G determined in the laboratory (accurate to slightly better than one part in 10^4)² and the value of G measured over distances of the order of one kilometre may have been detected by Stacey and his coworkers in analyses of variations in the acceleration of the Earth's gravity, g , in deep mines^{3,4}. Their results are not conclusive because the density of the attracting mass (the rock layers surrounding the mines) is not adequately known and is certainly inhomogeneous, and the possibility of regional gravity anomalies cannot be eliminated. Measurements of g in profiles through the ocean have also been suggested⁵. Although the density of seawater varies because of internal waves and small-scale mixing

We postulate that the presence of a relatively small number of VSV-infected cells can create widespread havoc either by cell-to-cell extension of the G antigen or by its propagation along the neuronal processes to considerable distances. The cytopathogenicity of VSV and the mode of extension of the lesions it induces may have parallels in both slow virus diseases and AIDS (acquired immune deficiency syndrome).

CHARLES CHANY
FRANÇOISE CHANY-FOURNIER
OLIVIER ROBAIN

Institut National de la Santé
et de la Recherche Médicale,
Hôpital St Vincent de Paul,
74 Avenue Denfert-Rochereau,
F-75014 Paris, France

1. Lifson, J.D. *et al. Nature* **323**, 725 (1986).
2. Sodroski, J., Goh, W.C., Rosen, C., Campbell, K. & Haseltine, W.A. *Nature* **322**, 470 (1986).
3. Chany, C. & Cook, M.K. *Annls Inst. Pasteur, Paris* **98**, 920 (1960).
4. Lepine, P., Chany, C., Droz, B. & Robbe-Fossat, F. *Ann. N.Y. Acad. Sci.* **81**, 62-72 (1959).
5. Chany-Fournier, G., Chany, C. & Lafay, F. *J. gen. Virol.* **34**, 305 (1977).
6. Handa, K., Chany-Fournier, F., Rousset, S. & Chany, C. *Biol. Cell* **44**, 261 (1982).
7. Florkiewicz, R.Z. & Rose, J.K. *Science* **225**, 721 (1984).
8. Robain, O., Chany-Fournier, F., Cerutti, I., Mazlo, M. & Chany, C. *Acta neuropath.* **70**, 220 (1986).

processes, it is certainly more uniform than that of rock. A major limitation in an oceanic measurement of G would be due to instrument motion and inadequate knowledge of the gravity meter's location within the water column.

Here we propose an alternative experiment using profiles of g obtained in boreholes through the polar ice cap. This offers several advantages over previously suggested approaches. The density of pure ice is controlled by temperature once the hydrostatic pressure is adequate to close cracks and squeeze out air inclusions⁶. Polar ice is remarkably free of impurities⁷, and we believe its density can be predicted from an equation of state with an accuracy approaching a few parts in 10^4 below depths of several hundred metres. This is likely to hold more strongly for the Antarctic than for the Greenland ice cap owing to the greater age of the ice and a reduced input of terrigenous contaminants. The density of polar ice is relatively homogeneous due to lateral flow of the ice sheet and uniform internal pressures from loading. There are also no difficulties from random, local accelerations of a gravimeter suspended in a borehole on the ice cap. Therefore, a polar ice borehole determination of G will have the instrumental stability and simplicity of the mine experiments along with the favourable density characteristics of an oceanic G determination.

There are currently at least two deep holes in the polar ice that are suitable for the proposed experiment. The first of these is located in south Greenland at Dye 3, and penetrates 2,037 m of ice to

bedrock in a region of rugged basement relief⁸. The second hole is at the Vostok Base in Antarctica and is over 2,000 m deep in an ice sheet of about 3,700 m thickness⁹. Both holes are filled with a hydrocarbon fluid having a density similar to that of ice to ensure that they do not close under hydrostatic pressure.

The experiment is envisaged as follows. A standard LaCoste–Romberg borehole gravity meter (LaCoste & Romberg Co., Austin, Texas) is lowered into the hole and a series of measurements of g at precisely determined depths is obtained. As this type of instrument measures g only in a relative sense and is subject to long-term drift, it must be calibrated elsewhere against an accurate absolute gravity meter¹⁰. Location in the hole can be determined by measuring the length of an invar wire under constant tension stretching from the instrument package to the surface. It can also be inferred from measurements of the ambient pressure in the borehole provided the density of the infilling liquid as a function of depth is known. More complex methods of depth measurement that could be used include surface radar tracking of a reflector on the borehole package or marine acoustic navigation techniques using a set of acoustic transponders on the ice interrogated in turn by one on the instrument package. In any case, an accuracy of one part in ten thousand should not be difficult to achieve. In practice, it will also be necessary to correct for gradients in gravity caused by topography at the ice–rock interface and variations in the thickness of the ice sheet. The form of such topography can be measured using precisely navigated ice-surface radar, a method in standard use by glaciologists¹¹. The major remaining source of systematic error is the presence of regional gravity anomalies caused by anomalous masses in the crust

and mantle underlying the ice. This argues strongly for performing the experiment in more than one borehole at different sites.

An experiment done over a 1 km vertical interval in which the uncertainties in ice density, gravity and depth within the borehole are respectively 1 part in 10^4 , $9\mu\text{Gal}$ ($9 \times 10^{-8} \text{ m s}^{-2}$), and 4 cm, will yield a value for G with an accuracy of 1 part in 10^4 . A tenfold degradation of this accuracy goal would still allow a meaningful comparison to be made with the value of G seen by others on this scale.

A.D. CHAVE

M.A. ZUMBERGE

Institute of Geophysics and Planetary Physics,

University of California, San Diego, La Jolla, California 92093, USA

M.E. ANDER

Earth and Space Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

J.A. HILDEBRAND

F.N. SPIESS

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA

1. Goldman, T., Hughes, R.J. & Nieto, M.N. *Phys. Lett. B* (in the press).
2. Luther, G.G. & Towler, W.R. *Phys. Rev. Lett.* **48**, 121–123 (1982).
3. Stacey, F.D., Tuck, G.J., Holding, S.C., Maher, A.R. & Morris, D. *Phys. Rev. D* **23**, 1683–1692 (1981).
4. Holding, S.C. & Tuck, G.J. *Nature* **307**, 714–716 (1984).
5. Stacey, F.D. *Geophys. Res. Lett.* **5**, 377–378 (1978).
6. Hobbs, P.V. *Ice Physics* (Clarendon Press, Oxford, 1974).
7. Hammer, C.U. et al. in *Greenland Ice Core: Geophysics, Geochemistry and the Environment* (eds Langway, C.C., Oeschger, H. & Dansgaard, W.) 90–94 (AGU, Washington DC, 1985).
8. Gundestrup, N.S. & Hansen, B.L. *J. Glaciol.* **30**, 282–288 (1984).
9. Kudriashov, B.B., Chistiakov, V.K., Zagrivniy, E.A. & Lipenkov, V. *Proc. 2nd International Workshop/Symp. Ice Drilling Technol.*, 123–124 (1982).
10. Zumberge, M.A., Rinker, R.L. & Faller, J.E. *Metrologia* **18**, 145–152 (1982).
11. Overgaard, S. & Gundestrup, N.S. in *Greenland Ice Core: Geophysics, Geochemistry and the Environment* (eds Langway, C.C., Oeschger, H. & Dansgaard, W.) 49–56 (AGU, Washington DC, 1985).

Mendel's legacy remains untarnished

SIR—Van Valen¹ suggests that I share the widespread view that Mendel falsified some of his data². In fact, I do not, though it is certainly the view of sources quoted by the author whose book I was reviewing³. I plead guilty to having tried to save space but not to iconoclasm. It is doubtful, however, if the accusations against the other “historic figures” (Burt and Dawson/Woodward) with whom I misleadingly coupled the name of Mendel would “also dissolve if similarly scrutinized”.

C.R.B. JOYCE

CIBA-GEIGY AG, CH-4002 Basel, Switzerland

Are ‘pragmatypes’ acceptable?

SIR—The rules for the designation of type specimens regularize the procedure for establishing the acceptable names for species. They minimize possibilities for dispute following advances in taxonomic understanding. For example, when a ‘species’ is found to be two species, which had been confused with each other, then the segregate which includes the holotype retains the existing name. If no holotype was designated when a species was originally described, a sub-sequent taxonomist may designate a specimen from its type series as a lecto-type. When the type material has been lost a neotype may be designated.

A notable advance this century has been the recognition of sibling species complexes. In many cases, internal features are the primary characters by which

such species have been recognized. Thus, cytotaxonomy has employed banding patterns in the giant chromosomes in dipteran salivary glands, and electro-phoresis has been used to highlight species-specific differences in enzymes. In such cases, the value of an old holotype specimen, impaled on a corroding nineteenth century pin, is minimal. Indeed, its state is likely to be an impediment to sensible decisions regarding the acceptable names for newly recognized species that had previously all carried the name attached to this holotype. This sort of experience caused Mattingly¹ to coin the term ‘mommet’ for such holotypes.

Recently, examination of sibling species–polymorphism complexes in Phoridae (scuttle flies) has led to recognition of the taxonomic value of features of the gut. These include an unusual crop mechanism² and the rectal papilla number, the external correlates of which have not always been discovered or in some cases exhibit an overlapping range. Thus most useful reference specimens are now the slide-mounts of dissected flies in which all the taxonomically familiar morphological features are displayed along with a display of the rectum.

Sorting specimens on rectal papilla number in the *Megaselia pulicaria* complex indicates that there are at least three species: A, B and C. The pinned holotype of *M. pulicaria* (Fallén) has been examined. Its rectal papilla number cannot be determined. What little of the external morphology can be inspected only allows an opinion that it probably belongs to species A.

It would seem sensible to designate a slide-mounted specimen of species A as a replacement type for *M. pulicaria*. This procedure would make Fallén's holotype of *M. pulicaria* a specimen of historical and curiosity value only. The primary reference specimen for the species, for nomenclatural purposes, therefore becomes the replacement type.

I should like to suggest the term ‘pragmatype’ for such replacement types. Clearly the designation of pragmatypes should only be countenanced when the primary characters used to separate newly recognized sibling species are features not observable in existing types. I intend to designate pragmatypes shortly, in a paper in preparation on the *M. pulicaria* complex, unless sound reasons are advanced against the proposal.

R.H.L. DISNEY

Department of Zoology, Cambridge University, Downing Street, Cambridge CB2 3EJ, UK

1. Van Valen, L.M. *Nature* **325**, 395 (1978).
2. Joyce, C.R.B. *Nature* **324**, 181 (1986).
3. Kohn, A. *False Prophets: Fraud and Error in Science and Medicine* (Blackwell, Oxford, 1986).

1. Mattingly, P.F. in *Taxonomy and Geography* (ed. Nichols, D.) (Systematics Association Publication No. 4, London, 1962).
2. Disney, R.H.L. *J. nat. Hist.* (in the press).

Search for natural isotope superlattices

SIR—The search for the formation of superlattice-type structures has yielded a good harvest for condensed matter scientists. Period doublings, commensurate-incommensurate transitions, spontaneous ordering of voids in metals under neutron irradiation, and so forth, all represent examples of self-organization (order-disorder transitions) in non-equilibrium systems in the sense of Prigogine synergetics¹. In all cases of self-ordering two requirements are important: first, there should be some interaction between ordering species distinct from the host interactions; second, the ordered state of the minority species (voids among metal atoms, electrons forming Wigner lattice, and so on) should correspond to the minimum energy configuration.

Here I would like to indicate another probable, though not yet observed, candidate for the superlattice club — namely, the isotope superlattices (ISL). The interaction between atoms of different isotopes is not exactly the same². Some small (about 1 part in 1,000) differences in lattice constants of isotopically pure crystals (for example, ²⁸Si and ³⁰Si) arise because of the anharmonicity of zero-point vibrations. In an isotopically mixed crystal this could lead to a 'net' isotope interaction between atoms of the same isotope that is roughly the difference between the interaction of different isotope pairs. This will not always lead to the ordering tendency — it will depend on the particular isotope ratio. But in many cases the ordered arrangement of the minority isotope will be energetically preferable.

It is, of course, highly improbable that any natural isotope ratios will be such that the formation of ISL will occur without stoichiometric mismatch. Also, the required kinetics of crystallization should fall into quite specific margins. This probably explains why ISL have not yet been detected in natural samples. However the required isotope ratios can easily be obtained experimentally. Various solidifying liquids, such as mixtures of light and heavy water, combined with various external conditions, such as irradiation or external fields, provide an enormous variety of experimental opportunities. Although interesting in its own right, the search for isotope correlations and ISL (both natural and artificial) could also have many applications. I have recently discussed³ the possibility of information storage based on isotope combinations.

ALEXANDER A. BEREZIN

Department of Engineering Physics,
McMaster University, Hamilton,
Ontario, Canada, L8S 4M1

1. Woo, C. H. & Frank, W. J. *nucl. Mater.* **137**, 7–21 (1985).
2. Berezin, A. A. *Chem. Phys. Lett.* **110**, 385–387 (1984).
3. Berezin, A. A. *Kybernetes* **15**, 15–18 (1986).

Amorphous definitions and glasses

SIR—In their recent paper Phillips *et al.*¹ present evidence for the existence of supernanometric microcrystallites in 'amorphous' silicon. We should admit, however, that there are many ways in which a solid does not have to be a crystal: the random network model of the structure of glasses does not always give way to clusters of microcrystallites.

There is incontrovertible evidence that quartz glass has a coherence length of about 1 nm, as expected for a random network: the thermal phonon² mean free path at 80 K is only 1.2 nm, as compared with 70 nm in quartz crystal (parallel to the *c* axis). There is no elastic scattering process known to me that can reduce a phonon to such a short path except phase cancellation between atomic motions on adjacent arms of a random network, by which is meant that a phonon moving from atom A to atom E proceeds in part by a long path ABDE and in part by a short path ACE. Between room temperature and 100 °C the mean free path in the glass is about 0.8 nm, consistent with structural coherence

of the same dimension. A structure of 3-nm microcrystallites would give proportionately longer mean free paths near room temperature, and the increase at low temperatures as we enter the Rayleigh region would be deferred to still lower temperatures.

The enormous photon extinction length of transmission-quality optical fibres is consistent with a random network model of the glasses used in the fibres. Rayleigh scattering³ in a wide frequency region is the dominant intrinsic mechanism that limits the extinction, and here we are talking about propagation over 100 km. Rayleigh scattering is proportional to the coherence volume: a change from (0.8 nm)³ to (3 nm)³ would have a measurable deleterious effect.

CHARLES KITTEL

Department of Physics,
University of California,
Berkeley, California 94720,
USA

1. Phillips, J.C., Bean, J.C., Wilson, B.A. & Ourmazd, A. *Nature* **325**, 121–125 (1987).
2. Kittel, C. *Phys. Rev.* **75**, 972–974 (1949).
3. Landau, L. & Lifshitz, E. *Electrodynamics of Continuous Media* Ch. 14 (Pergamon, Oxford, 1960).

Lung surfactant anchor

SIR—Patthy's¹ recent correspondence regarding the nature of the protein component of pulmonary surfactant is intriguing. As he points out, the major protein in both human and canine lung surfactant seems to have a collagen-like region within its amino-terminal half and a non-collagenous globular domain within its carboxy-terminal moiety. Of particular interest is the fact that the non-collagenous domain strongly resembles a module that is present in the proteoglycan of rat chondrosarcoma, the humoral lectin of the insect, *Sarcophaga peregrina*, and the membrane-associated lectins of human, rat and chicken hepatocytes.

In a recent cytochemical study of human neonatal lungs², I found that the filamentous surface coating of the epithelial cells ('pneumocytes') in the distal air spaces had binding sites for several lectins suggesting that the following carbohydrate groups are present in the filamentous material: N-acetylgalactosamine, N-acetylglucosamine, α-D-mannose, β-D-galactose and sialic acid. Another study³ indicates that the surface coating of hamster pneumocytes and alveolar macrophages also contains a similar range of terminal carbohydrate groups.

It is generally assumed that pulmonary surfactant is closely associated with the surface of the respiratory epithelium. The popular view, originally advanced by Pattle⁴, is that surfactant molecules form a monolayer at the interface between the alveolar air and the thin layer of fluid that moistens the epithelial cells. As the lectin

component of lung surfactant is similar to that derived from hepatocyte membranes and cartilage proteoglycans, it can be expected to be capable of binding to the terminal N-acetylglucosamine groups in the surface coating of the alveolar pneumocytes. The extension of the filaments into the fluid subphase of the surfactant monolayer would facilitate the binding process.

Patthy¹ has raised the possibility that the lectin-like component of surfactant protein interacts with the carbohydrate constituents of the lung surfactant complex. I share this opinion and, further, suggest that carbohydrate-lectin binding may help anchor the delicate surfactant film to the surface of the respiratory epithelium.

COWAN MEBAN

Department of Anatomy,
The Queen's University of Belfast,
Belfast BT9 7BL,
Northern Ireland

1. Patthy, L. *Nature* **325**, 490 (1987).
2. Meban, C. *Histochem. J.* **18**, 196–202 (1986).
3. Meban, C. *J. Anat.* **146**, 131–139 (1986).
4. Pattle, R. E. *Nature* **175**, 1125–1126 (1955).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

The stamp of greatness

W.F. Bynum

The Correspondence of Charles Darwin, Vol. 2 1837–1843. Edited by Frederick Burkhardt and Sydney Smith. *Cambridge University Press:1987. Pp.603. £30, \$37.50.*
The Darwinian Heritage. Edited by David Kohn. *Princeton University Press:1986. Pp.1,138. \$95, £63.40.*

THE second volume of the Darwin correspondence has now arrived, some months after the advertisements promised, but produced with the same devoted care as the first. As before, the notes helpfully explicate the letters, an appendix with manuscript variations and alterations should satisfy the most ardent Darwin pedant, and the biographical register and bibliography of books and articles mentioned make this a veritable reference guide to early Victorian natural science. The editing is by teamwork, and the product does much to restore one's faith in the committee as a means of producing works of scholarship.

The first volume saw Darwin safely through the *Beagle* voyage. This instalment finds him establishing himself in London, before marriage and the lure of the countryside's peace and quiet prompted him and Emma to settle in what was henceforth to be their headquarters, Down House. ("House ugly, looks neither old nor new", was Darwin's initial reaction.) By the end of 1843, Darwin was an established naturalist, one of England's scientific élite, a father and, despite his 34 years, something of an old man.

It is one of the bonuses of these early volumes of the correspondence that they cover sufficiently large blocks of time to enable us to witness developments in Darwin's preoccupations, style and psychology. The contrast between the Darwin of Cambridge and the *Beagle*, and the Darwin of London and Down House, could hardly be more striking. In the present volume, Darwin ages almost before our eyes. A series of short letters to Charles Babbage are entirely given over to accepting or declining invitations to his famous parties. Darwin manages to get up the energy to go to some of these soirées, usually with a relative or house guest in tow. As often as not, however, he declines, including one splendidly frank occasion in 1838 when he writes,

I am very much obliged to you for sending me cards for your parties, but I am afraid of accepting them, for I should meet some people there, to whom I have sworn by all the saints in Heaven, I never go out, & should, therefore, be ashamed to meet them.

Darwin's health will continue to fascinate historians, but although the complete correspondence is going to yield no definitive diagnosis, it already shows his health's

relatively rapid deterioration in the late 1830s. We know from previously published correspondence that it will be a recurrent theme of the remaining volumes.

This, however, is a trivial topic compared with the substance of Darwin's scientific achievement. Appropriately, the analytical index gives one column to his health, and more than three to his 'information networks'. Ill health was simply a nuisance, for the real business of life was

the advantage, of course, of having seen much of the South America that Darwin describes; nevertheless, his evaluation of Darwin's early achievement leaves the impression that Humboldt could see more clearly than any of Darwin's English contemporaries that this young naturalist was stamped not simply for distinction, but for greatness.

Humboldt died, aged 90, in that *annus mirabilis*, 1859. We now suspect that in January 1837 Darwin did not consciously believe in biological transmutation, whereas by December 1843 the basic structure that was to become the *Origin of Species* had been clearly worked out. This volume thus covers the crucial period of his scientific development. The notes to the letters occasionally point out themes and snippets of information which found their way into Darwin's notebooks, and the letters as a whole certainly make clear

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Darwin's home, Down House — "... ugly, looks neither old nor new", was his initial reaction.

discovery, and these letters reveal how single-minded and ambitious the young-old Darwin was. He took infinite pains to make certain that the thousands of specimens of plants, animals, rocks and fossils that he had collected were being properly described by the best people. He slaved over his *Journal of Researches*, fretted over his theory of coral reefs, agonized over his paper on the 'parallel roads' of Glen Roy, and engaged in lengthy exchanges with Lyell, Henslow, Waterhouse, Hooker, Owen, Gould, Herbert and many others on a vast array of topics. His uncle Joshua Wedgwood was right: Darwin was supremely "a man of enlarged curiosity".

Given Darwin's concern for his place within the community of scientists, the epistolary climax for him during these years (apart, he undoubtedly would have insisted, from letters from Emma) must surely have been what for us is also the most remarkable item of the collection: the long letter of 18 September 1839 from Alexander von Humboldt evaluating Darwin's *Journal of Researches*. Writing in French about a book in English, this great German naturalist and explorer recognized in Darwin a kindred spirit. He had

Darwin's intense concern, for reasons we now can understand, with animal and plant breeding. On balance, however, the correspondence reveals remarkably little of that private dialogue between himself and his notebooks. Darwin kept his own counsel for a very long time.

The various transmutation notebooks for this period have been published in several places over the past quarter of a century, but are shortly to be reissued in a single volume. This will be a great boon to Darwin scholarship, as is clear from *The Darwinian Heritage* in which the notebooks loom large. Few scientists besides Darwin would have secured such a large international group of scholars willing to contribute; even fewer might have found a publisher willing to produce a collection of articles which runs to over 1,100 pages. The 31 essays and commentaries cover most periods and themes relating to Darwin and darwinism; they can be taken as a fair reflection of the current state of Darwin scholarship, further examples of which may be found in the 70-page bibliography.

So much has been published on Darwin of late that new ideas can become scarce and 'scholarship' can be mistaken for the

academic equivalent of ghetto economics (everyone earning a living by taking in each other's dirty laundry, or commenting on each other's work, as the case may be). Only Darwin enthusiasts or compulsive reviewers are likely to read all of the essays in *The Darwinian Heritage*, but those who neglect the volume entirely will miss some sound research and fine writing. My own collection would be half the size, but it would certainly include Sloan on Darwin's invertebrate research, Browne on blushing, Kottler on natural selection, Rachootin on bones, Secord on

breeding and Mayr on Darwin's 'five theories'.

In the end, however, the essays are secondary literature, the correspondence is primary. Reading Darwin's correspondence may be more illuminating than reading about him. And on that scale, Darwin's scientific works stand even higher. Like *Middlemarch*, *The Origin of Species* will repay reading each year, from beginning to end. □

W.F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

Symmetrical work

Paul K. Townsend

Supersymmetry: A Decade of Development. Edited by P.C. West. *Adam Hilger*: 1986. Pp.483. £25, \$55.

Supersymmetry, Superfields and Supergravity: An Introduction. By Prem P. Srivastava. *Adam Hilger*: 1986. Pp.162. Hbk £25, \$55; pbk £12.50, \$28.

Introduction to Supersymmetry and Supergravity. By P. West. *World Scientific*: 1986. Pp.289. Hbk £31.80, \$36.80; pbk £18.90, \$21.85.

Introduction to Supersymmetry. By Peter G.O. Freund. *Cambridge University Press*: 1986. Pp.152. £20, \$34.50.

Unification and Supersymmetry: The Frontiers of Quark-Lepton Physics. By Rabindra N. Mohapatra. *Springer-Verlag*: 1986. Pp.309. DM84, \$34.

CONSIDERATIONS of symmetry are central to much of modern theoretical physics; certainly, symmetry was the guiding principle behind the achievements of elementary particle physics in the 1960s. In the early 1970s it was realized that quantum field theories may possess a new and radically different kind of symmetry — 'supersymmetry'. After an initial phase in which it was seen as something of a mathematical curiosity, supersymmetry caught the imagination of increasing numbers of theoretical physicists and underwent rapid development. In 1976 a supersymmetric generalization of general relativity, 'supergravity', was found. This gave rise to new hopes of a unification of gravitation and elementary particle physics.

Ten years and many man-hours later these hopes are still alive but in the form of 'superstring' theory. The spotlight on supergravity has passed on and it is now time to write a book about it. This is the premise of at least one of these five books on the subject. *Supersymmetry: A Decade of Development* is a collection of invited essays by means of which the editor has tried to achieve a "balanced view" of the subject. The first contribution, by

Golfand and Likhtman, can be recommended to those interested in the history of supersymmetry. These authors were early pioneers but their original paper (published in 1971) is hard to read, their methods being very different from those that followed, so the inclusion of this explanatory essay is all the more welcome.

The developments of the following "decade" (really 15 years) were prompted not by experimental support for supersymmetry but by the combination of its aesthetic appeal and the sheer difficulty of solving certain theoretical problems without invoking supersymmetric principles. In fact, so severe are the constraints imposed by supersymmetry that one of the considerable achievements of the period was the construction of supersymmetric theories that are not actually in conflict with experiment. These realistic models, as reviewed in the essay by G. Ross, do yield interesting new predictions but we shall have to wait for the next generation of particle accelerators to test them. Meanwhile, supersymmetry continues to find applications in other fields, such as cosmology and mathematics, the subjects of the two concluding essays. The other contributions are mostly expositions of the technical tools of the trade, many of them well-written. This is an excellent book that physics libraries would do well to acquire but which, by its nature, is not so suitable as a text for the individual student.

Because of its length, organization and choice of topics, P.P. Srivastava's *Supersymmetry, Superfields and Supergravity: An Introduction* invites comparison with *Supersymmetry and Supergravity* by Julius Wess and Jonathan Bagger (Princeton University Press, 1983). Both cover similar material on supersymmetry, but differ in their approaches to supergravity. Srivastava opts for the 'component' approach which is less geometrical than the superspace approach of Wess and Bagger but is more easily grasped on a first encounter and more easily explained. Even so, the 18 pages allotted are insufficient to do justice to the subject and this book really has to be considered as an introduction to supersymmetry alone. As

such it does a reasonable job, with plenty of problems for the student at the end of each chapter.

A more detailed and comprehensive account of supergravity is to be found in *Introduction to Supersymmetry and Supergravity* by P. West, which is about a hundred pages longer. The organization is also different in that both supersymmetry and supergravity are first discussed to some depth in the 'component' approach: only then is superspace introduced and the superfield formulation presented. For supergravity, at least, this is the historical route. Whether or not it is the best one, the book is the only one currently available that deals with both formulations adequately. Some material of more recent interest, on two-dimensional models and string theory, is included in the final chapters, but these are subjects that are rapidly evolving and have already moved on to the point that the chapter on strings, in particular, seems premature.

P.G.O. Freund's *Introduction to Supersymmetry* is refreshingly unconventional. Many traditional topics are left out (super-Feynman rules for example) and the treatment of supergravity is skimpy, but to compensate there is discussion of many other things such as Lie superalgebras and eleven-dimensional supergravity. This is also the only one of the volumes under review to present some of the mathematics needed for a rigorous definition of superspace. There is an entire book on this subject (Bryce DeWitt's *Supermanifolds*, published by Cambridge University Press in 1984), but Freund's brief discussion is likely to satisfy the needs of most physicists. His book can be recommended to those looking for a brief and not overly technical exposition of supersymmetry.

Unlike the other four books, *Unification and Supersymmetry* by R.N. Mohapatra is not a text on supersymmetry *per se*. It is rather an attempt to chart the course of theoretical elementary particle physics since the electroweak unification of the late 1960s. The book starts with an exposition of the basic ideas behind the electroweak theory and proceeds, for about two-thirds of its length, through topics such as CP-violation, 'grand-unification' and 'technicolour'. There follows an account of the applications of supersymmetry and supergravity to particle physics models. Despite its brevity, the choice of material here is sufficiently different to make this a useful book on supersymmetry. It is only a pity that the author does not discuss the more recent supergravity models that are currently favoured as 'low-energy' manifestations of superstring theories. □

Paul K. Townsend is a Lecturer in the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

Black holes made easy

William H. Press

Black Holes: The Membrane Paradigm. Edited by Kip S. Thorne, Richard H. Price and Douglas A. Macdonald. *Yale University Press: 1986. Pp. 367. Hbk \$37.50, £37.50; pbk \$12.95, £9.95.*

WHERE are the black holes? Have they gone the way of steady-state cosmology and phlogiston, or are they still an essential element of contemporary astrophysical understanding?

The answer is that black holes, of about 10^8 solar masses, are alive and well and living in quasars and active galactic nuclei. This blunt statement is accepted by most researchers in the field, but it has been unexpectedly difficult to prove by any single, direct observation. Rather, the past decade has seen the continuing accumulation of data and theoretical understanding consistent with the black-hole model, and not very consistent with any other model. These advances may no longer have the ability to grab popular headlines (or stampedes of PhD students), but steady progress is punctuated by elegant new insights and syntheses. One of the most fruitful of them is the 'membrane viewpoint', developed by Kip S. Thorne and his collaborators, several of whom are contributors to this book.

The idea is this: Suppose we are willing to restrict our study to the interaction of already formed black holes with the external, astrophysical universe, putting aside for another book questions of black-hole formation and interiors. Then we want to think of the black hole not as a 'spacetime', but as an 'object', a three-dimensional approximately spherical 'thing' that behaves like any other astronomical body — accretes material, interacts with external matter and electromagnetic fields, and so forth. The surface of this object can be taken to be a timelike surface, a membrane, just outside (in a sense that can be made mathematically precise) the actual event horizon of the black-hole spacetime. The game now is to 'hide' the inside of the black hole by attributing all its properties to fictional, but precisely definable, properties of the membrane.

Physically motivated thought-experiments are conducted to investigate the properties of the membrane, its electrical conductivity for example. This can be discerned in several ways, say by calculating the membrane's absorption of electromagnetic radiation, or its Ohmic losses in a time-varying magnetic field, or from the details of how the electric field on the membrane becomes uniform if a charge is dropped onto (into!) one point. Amazing-

ly, all such calculations give the same answer for the conductivity. The deep reason is that the exact, fully relativistic equations for electromagnetism in a black-hole spacetime can be cast into a form *exactly* like ordinary Maxwell's equations in the presence of a conducting body.

This general idea works not only for the hole's electrical properties, but for its mechanical and thermodynamic properties as well. The membrane has a viscosity that gives rise to dissipation whenever an orbiting object raises a tidal bulge — just like the Moon's tidal dissipation in the oceans of the Earth. The membrane has an entropy and a temperature, in terms of which the black hole and its surroundings can be understood as exactly satisfying the first and second laws of thermodynamics. Once again, the analogies are not merely suggestive, they are *exact* when the equations are properly cast. The membrane viewpoint is a powerful formalism, not only for actual calculations of black-hole behaviour, but also for intuitive and qualitative understanding of a hole's behaviour in the complex astrophysical

environment at the centre of a quasar or active galactic nucleus.

Although the chapters in this book were written by nine separate authors, they have been polished to speak with a single voice: it is the voice of a careful teacher who, while unafraid of mathematical precision, is not satisfied with an equation until a clear physical interpretation of it has been conveyed. By virtue of its pedagogical excellence, the book transcends its particular (membrane) viewpoint and is, quite simply, about the best single work on black holes yet written. The illustrations are numerous and clear, and the index is superb. Above all, it is a book in which physics never takes a back seat to formalism. Backed up by Chandrasekhar's uncompromisingly technical *The Mathematical Theory of Black Holes* (Oxford University Press, 1983), it represents the current and maturing state of knowledge about these exotic — but not necessarily unintuitive — astronomical bodies. □

William H. Press is Professor of Astronomy and Physics at Harvard University, Cambridge, Massachusetts 02138, USA.

Of barriers and carriers

S.R. Caplan

Principles and Models of Biological Transport. By Morton H. Friedman. *Springer-Verlag: 1986. Pp. 260. DM 165.*

Books on transport, especially biological transport, are two-a-penny (figuratively speaking) and at first glance it is not easy to discern why we need yet another treatise on the topic. Can there be any reason to repeat what has only just been said with great profundity in 685 pages, but this time in less than half the space?

It seems that there may be. Brevity has its advantages, especially where wordiness is replaced by precise description in mathematical terms. Because this monograph is intended to provide an introductory course for undergraduate biomedical engineers and graduate students in biophysics, the author can afford to present most of the material as an exact science, highlighting the main concepts without becoming bogged down in a mass of experimental minutiae. In this way a remarkably wide variety of topics is covered, including thermodynamics of transport processes (both the equilibrium and nonequilibrium varieties), diffusion (free and facilitated), active transport, kinetic and other models, single-cell and epithelial phenomena, gaseous processes and much else besides. Nerve, muscle and kidney all receive due attention.

Unfortunately, several caveats must be recorded. The fact that difficulties are glossed over (and important issues ignored) is less bothersome than the author's tendency to trot out often-repeated misunderstandings. Among these is the claim that the "dephosphorylation reaction" (of ATP) "is accompanied by the release of the energy stored in the terminal phosphate bond" (p.78). This actually precedes by a few pages a prominent reference to a paper by Eisenberg and Hill, wherein the surprised reader will find this view soundly debunked. Elsewhere the condition generally known as detailed balancing is said not to hold for active transport (p.101), and the reflection coefficient is indicated to lie between zero and unity (p.119).

The sternest caveat, however, relates to the inadequate and arbitrary referencing of sources. Quoting from contemporaries without citation is a serious discourtesy, which, moreover, denies the reader access to the most relevant literature. By contrast, the gentleman cited repeatedly as "van t'Hoff" might have welcomed anonymity!

Still, this is a book that will repay study by those who keep their wits about them. □

S.R. Caplan is a Professor in the Department of Membrane Research, The Weizmann Institute of Science, 76 100 Rehovot, Israel.

•In his review last week (p.218), Robert Whittle expressed the hope that Wiley would issue a "more imaginatively" priced edition of *Genetic Analysis of Animal Development*. The book will appear in paperback in May; price will be £25.40, \$34.95, against £66.95, \$69.95 for the hardback.

Birdlife down on the farm

John Andrews

Farming and Birds. By Raymond J. O'Connor and Michael Shrubbs. Cambridge University Press: 1986. Pp.290. £17.50, \$34.50.

FARMING is the main agent of the destruction of terrestrial wildlife habitats throughout the world. Sometimes the process is quick and obvious, such as the clearance of forest or the drainage of wetland. Often it is gradual and imperceptible, for instance where domestic livestock introduced to natural grassland graze-out the palatable vegetation. At the same time, there have always been species with adaptations suited to the new circumstances and with the powers of dispersal necessary for colonization. Some birds have altered their world ranges in response to opportunities created by agriculture, many have increased in abundance and a few become economically significant pests. Where the entire land surface is man-modified, farms may retain semi-natural areas which are little used and remain relatively species-rich, while the farming system itself may sustain most

or all of the populations of some species — grey partridge, rook and whitethroat in Western Europe, for example.

In both Europe and the United States the future of agriculture is unclear, mainly because of over-production, and conservation bodies that wish to influence that future must understand and be able to demonstrate the ecological effects of different farming practices. The publication of *Farming and Birds* is therefore most timely, particularly because it is, perhaps surprisingly, the first systematic attempt to review knowledge of the impact of British agriculture on the country's birdlife. It confirms the subjective view of many naturalists that recent changes in practice have been generally detrimental to birds and hoists a warning flag over the future of several species. Moreover, its broad findings probably have relevance in all countries with technically advanced agriculture.

The senior author, Raymond O'Connor, is director of the British Trust for Ornithology (BTO) which has built up data on bird population changes for over 25 years, mostly in the form of systematic observations made by hundreds of skilled amateurs. Michael Shrubbs is one such amateur and, more importantly, a farmer with a wide understanding of agricultural systems. They have drawn on four sources of data — the BTO's Common Birds Census and Nest Record Scheme, including much previously unanalysed material; the existing literature, some of it unpublished; the Ministry of Agriculture's annual farm census returns which show changes in cropping patterns; and surveys of fertilizer usage. The volume of data available was such that a 1980 cut-off date was applied.

Farming is a dynamic activity and the authors commence by summarizing its history from the Middle Ages to the Second World War, showing how patterns of grass cover and tillage change in response to economic forces. Next they describe Britain's farmland bird community, which has always been small in terms of numbers of species. The original climax vegetation being woodland, Britain has few birds adapted to life in open fields, unlike North America for instance. At the core of the farm avifauna are woodland birds which use hedges, farm woods and farmstead gardens for nesting and for much of their foraging. Their overall distribution is determined by climate, soil type and terrain, both directly and indirectly through the influence of these factors on farming procedures. The value of farmland as nesting and foraging habitat is assessed in relation to performance in primary habitats and in different farming contexts. The authors conclude that mixed farming systems, combining tillage crops and livestock within holdings averaging about 25 hectares, are most favourable to bird diversity and abundance.

Mechanization and chemical technology increasingly permit specialization and intensification in agriculture, and the effects on birdlife are considerable. A number of them are considered, including the loss of feeding opportunities associated with the reduction in the number of mixed farms; the removal of features such as hedges and ponds; the intensification of use of semi-natural farmland (such as the conversion of old permanent pasture);

Eric Hosking

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Not sitting pretty — the lapwing (Vanellus vanellus), which is under threat from new farming practices in certain regions of Britain.

changes in the timing of cropping, including the shift to autumn-sown cereals; and the use of pesticides and artificial fertilizers. Several birds, common until recently, including linnet, grey partridge and lapwing, are seen as being at risk of extinction in certain regions.

The book is open to criticism on several counts. There are some dangerous simplifications, such as the statement that "wildlife can adapt to change and has done so in the past but it cannot keep pace with the speed of modern technological development". A number of conclusions are tentative, because data are incomplete, or have been inferred from the fact that changes in bird numbers or distribution coincide with change in farming practice. And some studies by other workers appear to be accepted uncritically. The fact is, as the authors themselves state that knowledge of the effects of farming operations on birds is still very limited and the ecology of farmland species is understood only poorly. Nonetheless this is a valuable exercise, readable if somewhat repetitious, with many pointers for further study in Britain and elsewhere. □

John Andrews is Chief Advisory Officer for The Royal Society for the Protection of Birds, The Lodge, Sandy, Bedfordshire SG19 2DL, UK.

NATIONAL GEOGRAPHIC Research

a quarterly scientific
journal for scientists,
students, researchers. . .

Written and edited by scientists, it offers multidisciplinary coverage in fields such as anthropology, ecology, geography, and zoology. It features:

- original manuscripts, peer review
- technical detail, color graphics
- indexing by major services
- open submissions policy
- book reviews, letters, and more.

Just \$40*

for an annual subscription in the U.S.
Reader Service No.1

YES, I want to subscribe to NATIONAL GEOGRAPHIC RESEARCH with no obligation. Please send me an examination copy of the current issue and enroll me as a trial subscriber.

NAME

75109

STREET

CITY STATE/PROVINCE COUNTRY ZIP/POSTAL CODE

*Canada — \$45 (\$63.75 Canadian funds); all other countries — \$45 surface, \$60 airmail.

Mail to: National Geographic Society, P.O. Box 1111 Washington, D.C. 20013-1111 U.S.A.

Seismological measurement of stellar ages

Douglas Gough

Institute of Astronomy and Department of Applied Mathematics and Theoretical Physics, University of Cambridge, CB3 0HA, UK

Accurate measurements of stellar acoustic oscillation frequencies, which are likely to be obtained in the coming years, will provide important new data enabling us to diagnose the internal structure of stars. The ages of main-sequence stars are likely to be determined to a higher precision than has been possible in the past and much more accurate estimates of the total heavy-element abundances should be possible.

WE shall never have from the stars a wealth of seismic data such as has been gathered from the Sun¹; variations across the surface of a distant star cannot be well resolved, and therefore only those modes of oscillation with low degree l can be detected. (The spatial variation of the vertical velocity and optical intensity fluctuations associated with a mode of oscillation is proportional to a spherical harmonic of degree l .) An impression of what might be obtained is illustrated in Fig. 1, which is a temporal power spectrum of the line-of-sight (Doppler) frequency shift of the 769.9 nm K line in light that has been spatially integrated over the entire disk of the Sun, as would be light from the stars. Only modes with $l \leq 3$ are evident². Their discrete cyclic frequencies $\nu_{n,l}$ are approximated by the asymptotic formula³, valid when $n \gg l$:

$$\nu_{n,l} \sim (n + \frac{1}{2}l + \tilde{\alpha})\nu_0 + \varepsilon_{n,l} \quad (1)$$

where n is the order of the mode, $\tilde{\alpha}$ is a constant of order unity, and $\varepsilon_{n,l} = O(\nu_0 n^{-1})$ is small compared with $\nu_{n,l}$; in the figure $n' \equiv n + \frac{1}{2}l$ lies between 13.5 and 28. The constant ν_0 is a characteristic frequency of the star, defined as the harmonic mean with respect to radius r of the speed of sound $c(r)$ divided by the diameter of the star.

As indicated by equation (1), the peaks in the power spectrum are grouped in pairs, the groups being separated by a frequency interval $\Delta_n \approx \frac{1}{2}\nu_0$. Furthermore, modes of a group with like n' are separated in frequency by $d_{n,l} \equiv \varepsilon_{n,l} - \varepsilon_{n-1,l+2}$. With the theoretical knowledge that $d_{n,l} \propto 2l+3$ and that $d_{n,l} > 0$, it is possible to ascertain whether modes in a group have odd or even l , and then to identify the value of l pertaining to each of the peaks in the power spectrum.

The modes under discussion have substantial amplitudes throughout most of the interior of the star; they avoid merely

a central sphere, which for solar-like stars has radius $r_c \approx 0.8(l + \frac{1}{2})R/(n' + \tilde{\alpha})$, where R is the radius of the star. Thus some of them penetrate into the central core where nuclear reactions are taking place.

Acoustic modes can be regarded as a superposition of sound waves which propagate between the stellar surface and the radius r_c at which downward directed waves are refracted back towards the surface. When $n \gg l$ their frequencies form an approximately harmonic sequence which depends on the size of the acoustic cavity within which the waves propagate (from $r = r_c$ to $r = R$), and on the (harmonic) mean of the speed of sound c within that cavity. Thus, because $r_c \ll R$, the cavity is almost the entire star and therefore the frequency interval Δ_n depends only weakly on the value of r_c , and hence only weakly on n' ; it is approximately $\frac{1}{2}\nu_0$, as is evident from equation (1), which is a global average of c . The separation $d_{n,l}$ on the other hand, which measures the extent to which the overtones deviate from an harmonic sequence, is the difference between two similar averages of c and is therefore dominated by conditions between the two values of r_c associated with the modes whose frequencies are differenced. Thus $d_{n,l}$ depends most sensitively on conditions in the core.

As a star evolves on the main sequence, hydrogen is converted into helium in the core, which increases the mean 'molecular weight' of the stellar material near the centre of the star and decreases the sound speed. As a result of this variation alone, acoustic frequencies are reduced. The reduction is greater for modes with smaller l , which penetrate more deeply into the core. Therefore $d_{n,l}$ is reduced. The dominant response of the nonreacting envelope of the star is to expand. The size of the acoustic cavity is increased, and consequently Δ_n is reduced too.

The structure of a main-sequence star of given initial chemical composition depends only on its mass M and age t . Therefore, in principle, it should be possible to determine both these quantities from a knowledge merely of representative, though well defined values of Δ_n and $d_{n,l}$. Indeed, as a first approximation, zero-age lower main-sequence stars, for example, may be considered to form an homologous sequence satisfying $M \propto R^{3/4}$. Under homologous transformation the overtone frequency interval satisfies the 'pendulum formula': $\Delta_n \propto (g/R)^{1/2} = (GM/R^3)^{1/2}$, where g is the surface gravity, and is proportional to the square root of the mean density of the star. Therefore $\Delta_n \propto M^{-3/2}$, a result which is confirmed by numerical computation^{4,5}. As the star subsequently evolves its radius increases and the homology scaling continues to hold: $\Delta_n \approx 135(M/M_\odot)^{1/2}(R/R_\odot)^{-3/2} \mu\text{Hz}$. Thus Δ_n depends on both mass and age. Notice, however, that surface gravity g can be estimated spectroscopically, and together with Δ_n can be used to determine M and R separately. The separation $d_{n,l}$ measures the nonhomologous change in the core, and therefore determines principally what one might call the evolutionary age $\tau = t/t_{\text{ms}}$,

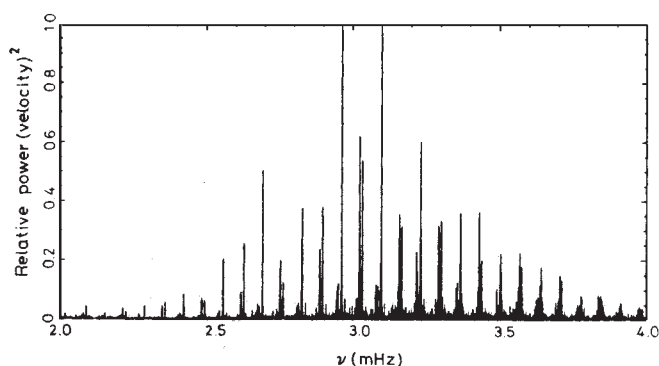


Fig. 1 Power spectrum of 88 days of nearly continuous observations of whole-disk spectrum line shifts obtained from stations in Tenerife and Hawaii, which are separated by 9.7 h in longitude (after ref. 17). The maximum amplitude = 15 cm s⁻¹.

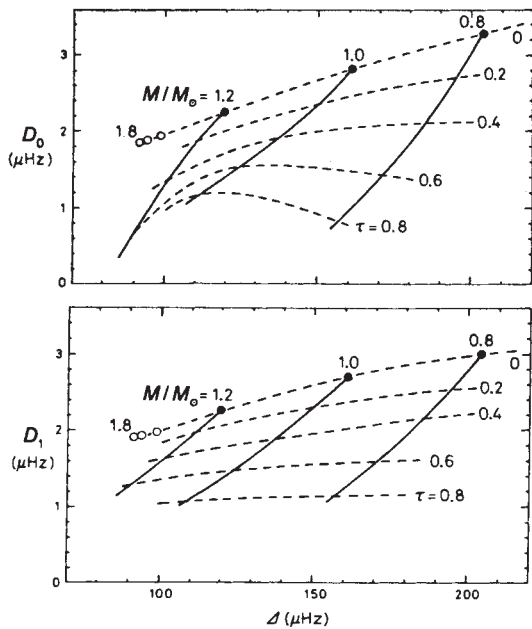


Fig. 2 Continuous lines are contours of constant M , dashed lines are contours of constant τ , estimated from theoretical data in the frequency range 1.5–3.8 mHz provided in refs 5 and 13 for main-sequence stellar models with $Y = 0.257$ and $Z = 0.017$. The dashed curves, in particular, were obtained by interpolation between points on the continuous curves, and in reality may vary more sharply than depicted in the bottom left-hand corner of the upper diagram where it is evident that the frequencies of the lowest-degree modes are influenced by the convective core that develops in stars somewhat more massive than the Sun. Physical time is $t = t_{\text{ms}}\tau$; the values of t_{ms} are 22.8, 9.1 and 4.3×10^9 yr for $M = 0.8$, 1.0 and $1.2 M_{\odot}$ respectively. The circles on the zero-age contours are at intervals of $0.2 M_{\odot}$ in mass; filled circles were obtained from refs 5 and 13, and open circles, which are representative values for both $l = 0$ and $l = 1$ in the frequency range 2–4 mHz, from ref. 4.

where I define t_{ms} to be twice the time taken for the central hydrogen abundance to be reduced to half its initial value. Of course t_{ms} depends on M , and for solar-type stars $t_{\text{ms}} \propto M^{-4}$. Hence, even though $d_{n,l}$ might be expected to depend relatively weakly on M at fixed τ , it is necessary to know both τ and M (and thus both $d_{n,l}$ and $\Delta_{n,l}$) to ascertain the absolute age t .

Seismic measurements of lower main-sequence stars have not yet attained adequate sensitivity for measuring $d_{n,l}$. Indeed, individual modes of oscillation have not even been isolated. Nevertheless, Fossat *et al.*^{6,7} and Noyes *et al.*⁸ have between them searched for oscillations in three late-type main-sequence stars, α Centauri A, Procyon and ϵ Eridani, and have found statistical evidence for modes whose frequencies satisfy equation (1). They have been able to estimate only ν_0 from their observations, and therefore the astrophysical inferences that have so far been made rely heavily on other astronomical evidence^{9–11}.

Isolating the modes in a power spectrum such as Fig. 1 requires long, almost continuous data sets; under ideal conditions it would be necessary to observe for about two weeks to attain a precision of $1 \mu\text{Hz}$, for example, in the measurement of $d_{n,l}$. Several groups of astronomers are planning such measurements, either by coordinated observations from two or more ground-based sites around the world, or by observing from space.

In anticipation of imminent progress, Christensen-Dalsgaard⁴ has carried out the first computations to quantify the theoretical ideas I have outlined above. In particular, he has plotted for several values of M at $t = 0$ and for several values of t at $M = M_{\odot}$ a representative value (which he has called D_0) of $D_{n,l} = d_{n,l}/(6 + 4l)$, which is asymptotically independent of l , against a representative value ($\Delta \nu_0$) of $\Delta_{n,l}$. Both quantities were obtained by fitting numerically computed eigenfrequencies to the

asymptotic formula (1). The result confirmed that, at least for stars with masses similar to that of the sun, contours of constant M and constant t are not even nearly parallel (see for example, Fig. 2), and therefore it appears that in principle it would indeed be possible to determine both M and t from $D_{n,l}$ and $\Delta_{n,l}$. More detailed results¹² were presented at the IAU Symposium 123 on Advances in Helio- and Asteroseismology held in Aarhus last July.

Results similar to those of Christensen-Dalsgaard are illustrated in Fig. 2, where contours of constant M and τ are plotted against representative values D_l and Δ of $D_{n,l}$ and $\Delta_{n,l}$. They were derived mainly from recent computations by Ulrich^{5,13}. It appears that in principle either one of the plots could be used to determine M and τ simultaneously from Δ and either D_0 or D_1 ; I present both diagrams to show how the seismic data contain even more information. In particular, the relatively high curvature of the isochrones in the bottom left-hand corner of the Δ – D_0 plane is associated with the development of convective cores in the more massive stars. It is important to realise, however, that the determination of M and τ could be accomplished by comparing observations with these diagrams only if the chemical composition, and the mixing-length parameter α used in modelling the convection zone, can also be ascertained. That requires the use of additional astronomical data.

The structure of theoretical stellar models depends on M , t , α and chemical composition, which is usually characterized by any two of the parameters X , Y and Z , the initial relative abundances by mass of respectively hydrogen, helium and all other elements ($X + Y + Z = 1$; the relative abundances of the heavy elements that influence stellar structure most strongly is usually assumed to be known¹⁴). For given X , Z and α , stellar evolution theory relates the luminosity L and the surface effective temperature T_e to M and t . Thus if L , T_e and say Z/X are determined astronomically, as might be expected for stars in clusters, the seismic parameters Δ and D_0 , say, are then formally sufficient to determine M , t , Y and the mixing-length parameter α . If the star is one of a binary pair, so that M is also known, one would then expect it to be no longer necessary to know the chemical composition. But how accurate would the determinations be?

A step towards answering that question has been made in two recent papers by Ulrich, which address separately the seismological determination of mass¹³ and age⁵. By combining the results from the two papers one can estimate the sensitivity of the determinations to errors in the data. An error matrix for a solar-type star is presented in Table 1, which lists the partial derivatives of t/t_{ms} , $\ln M$, $\ln Y$ and $\ln \alpha$ with respect to the parameters listed in the first column; the units of these parameters are stated in parentheses, and represent perhaps somewhat optimistic estimates of the precision with which the parameters can be measured. The most striking feature is the high sensitivity to the relatively poorly known heavy-element abundance Z , which, even ignoring any doubts one might have about assuming the value of Z in the stellar interior to be the same

Table 1 Formal errors in inferred age t , mass M , initial helium abundance Y and a mixing-length parameter α

	$\delta t/t_{\text{ms}}$	$\delta M/M$	$\delta Y/Y$	$\delta \alpha/\alpha$
T_e (2%)	–0.06	0.04	–0.09	0.16
L (10%)	0.03	0.05	–0.02	–0.11
Z/X (50%)	1.1	–1.9	5.6	–2.6
Δ (1 μHz)	0.01	0.01	–0.02	–0.003
d_0 (1 μHz)	–0.13	–0.02	0.10	–0.03

Formal errors for a solar-type star resulting separately from errors in the measurements of T_e , L , Z/X , Δ and d_0 (of the magnitudes stated in parentheses) computed from the data in refs 5 and 13. The parameter d_l is $(6 + 4l)D_l$ and is thus a representative value of $d_{n,l}$. The effect of replacing d_0 by d_l is to multiply each element of the last row by approximately 3/5; the remaining elements are hardly changed.

Table 2 Transformation of the error matrix of Table 1, with g replacing Z/X

	$\delta t/t_{\text{ms}}$	$\delta Z/Z$	$\delta Y/Y$	$\delta \alpha/\alpha$
T_e (2%)	0.04	0.07	0.12	-0.005
L (10%)	-0.03	-0.02	-0.24	-0.008
g (20%)	-0.17	-0.07	-0.57	0.28
Δ (1 μHz)	0.01	0.01	-0.002	-0.013
d_0 (1 μHz)	-0.14	-0.04	0.05	-0.003

as that in the atmosphere, renders it impossible to determine either M or t . This conclusion is at variance with the view advanced by Ulrich.

Use of additional, more subtle features of the frequency spectrum or independent measurements of other stellar parameters could permit the determination of t . In particular, the situation is eased substantially by using a separate estimate of M , either indirectly from the surface gravity g , or directly if the star is a member of a binary system. In this case, T_e and L determine R , so M/R^3 can be estimated directly and the role of Δ , like d_0 , is now solely to provide information about deviations from homologous variation. The outcome is illustrated in Table 2, which is the error matrix for the case when g is measured. In practice, spectroscopic measurements actually estimate combinations of quantities such as T_e , L , g , Z/X and Y , and it is a straightforward matter to transform the error matrix for any given self-consistent complete set of such combinations.

It is evident from Table 2 that one can now hope to know the age of a solar-type star to within perhaps 20% of its main-sequence lifetime t_{ms} . Moreover, one can also measure Z , which determines the opacity in the interior of the star. It would therefore be particularly interesting to make seismological observations of a metal-deficient late-type star; B. E. J. Pagel (personal communication) has suggested ν Indi, which is a fifth magnitude G star in the Southern Hemisphere. Unfortunately, according to Table 2, without additional data one cannot determine Y to useful accuracy.

It should be appreciated that the measurement of t and Z is based on a calibration of theoretical models, and therefore depends on the accuracy of those models. Therefore one must heed the fact that it has not yet been possible to reproduce accurately the seismic data from the Sun by adjusting the parameters of the theory; until we truly understand why that is so we must treat any calibration of stellar models with due caution. In the solar case, the slight, possibly insignificant discrepancy^{15,16} between the observed and calculated values of D_0 , for example, might imply that the Sun is somewhat older than the 4.7×10^9 yr that is generally presumed; alternatively, there may be some important physical process that has not been adequately described by the standard theory. Nevertheless, it is evident that any seismic data we do obtain from stars is likely to improve our knowledge of those stars considerably.

I thank M. G. Edmunds, J. W. Harvey, B. E. J. Pagel and M. J. Thompson for useful conversations.

- Christensen-Dalsgaard, J. Gough, D. O. & Toulme, J. *Science* **229**, 923-931 (1985).
- Christensen-Dalsgaard, J. & Gough, D. O. *Mon. Not. R. astr. Soc.* **198**, 141-171 (1982).
- Tassoul, M., *Astrophys. J. Suppl. Ser.* **43**, 469-490 (1980).
- Christensen-Dalsgaard, J. *Proc. Workshop on Space Research Prospects in Stellar Activity and Variability* (eds Mangeney, A. & Praderie, F.) (11-44 (Obs. Paris-Meudon, 1984).
- Ulrich, R. K. *Astrophys. J.* **306**, L37-L40 (1986).
- Fossat, E., Grec, G., Gelly, B. & Decanini, Y. *C. r. Acad. Sci. Paris, Ser. 2* **229**, 17-20 (1984).
- Fossat, E., Gelly, B. & Grec, G. *Seismology of the Sun and the Distant Stars*: 405-416 (ed. Gough, D. O. 1986).
- Noyes, R. W. *et al. Astrophys. J.* **285**, L23-L26 (1984).
- Gough, D. O. *Nature* **314**, 14-15 (1985).

- Demarque, P., Guenther, D. & van Altena, W. F. *Astrophys. J.* **300**, 773-778 (1986).
- Guenther, D. B. & Demarque, P. *Astrophys. J.* **301**, 207-212 (1986).
- Christensen-Dalsgaard, J. in *Advances in Helio- and Asteroseismology*, (eds Christensen-Dalsgaard, J. & Frandsen, S.) (Reidel, Dordrecht, in the press).
- Ulrich, R. K. in *Advances in Helio- and Asteroseismology* (eds Christensen-Dalsgaard, J. & Frandsen, S.) (Reidel, Dordrecht, in the press).
- Ross, J. E. & Aller, L. H. *Science* **191**, 1223-1229 (1976).
- Faulkner, J., Gough, D. O. & Vahia, M. N. *Nature* **321**, 226-229 (1986).
- Däppen, W., Gilliland, R. L. & Christensen-Dalsgaard, J. *Nature* **321**, 229-231 (1986).
- Claverie, A., Isaak, G. R., McLeod, C. P. & van der Raay, H. B. *Mem. Soc. astr. Italiana* **55**, 63-66 (1984).

ARTICLES

A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*

Edwin L. Ferguson, Paul W. Sternberg* & H. Robert Horvitz

Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Twenty-three genes have been assigned to particular steps in a genetic pathway for the specification of the vulval cell lineages of the nematode Caenorhabditis elegans. Mutations in most of these genes cause homoeotic transformations in the fates of individual cells, suggesting that these lineages may be specified by a series of decisions that distinguish between alternative cell fates. Fifteen of the genes function in a system involved in the intracellular response to the extracellular signal that induces vulval formation.

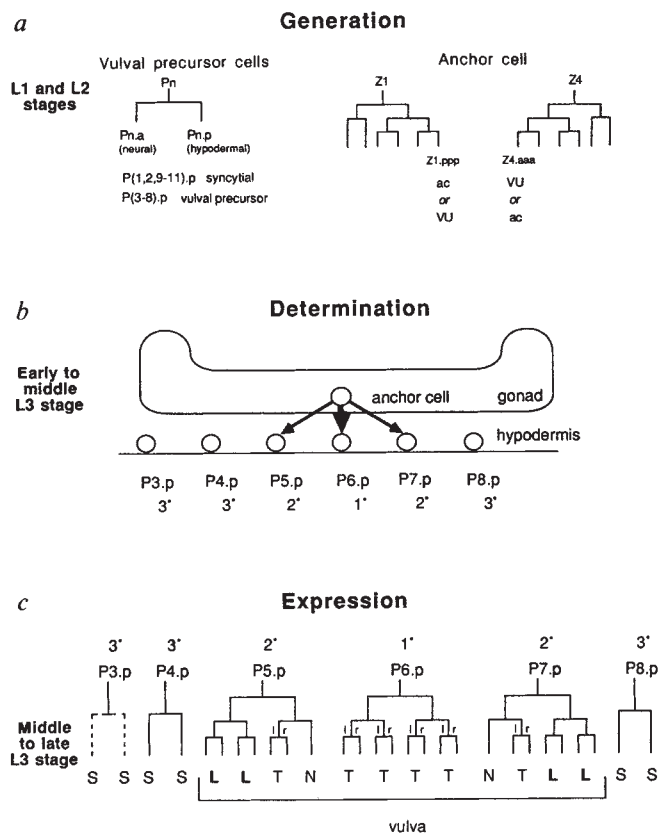
To understand how patterns of cell divisions and cell fates are specified during development, the complete cell lineage of the nematode *Caenorhabditis elegans* has been determined and extensively analysed¹⁻⁸. We have focused our attention on the cell lineages involved in vulval development for three main reasons. First, the vulval cell lineages comprise only three rounds of cell division that occur over a period of only five hours¹. Second, the specification of the vulval cell lineages appears to

involve both cell-autonomous and non-cell-autonomous decisions⁵⁻⁸. Third, certain mutants abnormal in the vulval cell lineages can be recognized easily with a dissecting microscope and are viable as homozygotes⁹⁻¹².

So far, we have identified more than 100 mutants that have one of two phenotypic abnormalities in vulval morphology. Multivulva (Muv) mutants have multiple vulva-like protrusions along their ventral sides; vulvaless (Vul) mutants lack a vulva. We believe that we have identified most genes for which lack of gene function results in a Muv or Vul phenotype (see ref. 12 for discussion). Most of the genes defined by Muv or Vul mutations have been named *lin*, for lineage abnormal.

* Present address: Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94143, USA

Fig. 1 Developmental pathway for the vulval cell lineages (based on refs 1, 5, 7 and 8). **a**, Generation of the seven cells involved in vulval development, P(3-8).p and the anchor cell (ac) (see text). In these diagrams, each vertical line represents a cell and each branch represents a cell division. Divisions are anterior-posterior, unless otherwise indicated; the anterior daughter is indicated by the left branch. The cells involved in vulval development are named according to their ancestries; for example, P6.p is the posterior daughter of P6, and Z4.aaa is the anterior daughter of Z4.aa, which is the anterior daughter of the anterior daughter of Z4. (See ref. 1 for a description of the conventions used in drawing cell lineages and in naming cells.) Pn, any of the 12 ectodermal progenitor cells P(1-12). Z1 and Z4, the two somatic gonadal progenitor cells. ac, the anchor cell. VU, a ventral uterine progenitor cell. **b**, Determination of the fates of the vulval precursor cells. During the early to middle L3 stage, a graded inductive signal (indicated by the arrows) from the gonadal anchor cell specifies which of three lineages (1°, 2°, or 3°) each of the six tripotent precursor cells P(3-8).p will express (see text). **c**, Expression of the vulval lineages. During the middle to late L3 stage, P(3-8).p express their respective lineages. P(3,4,8).p express a 3° lineage; these cells divide once and join the hypodermal syncytium. Each cell that joins the hypodermal syncytium is represented by [S], so the 3° lineage is represented by [S S]. (In ~50% of wild-type hermaphrodites, P3.p joins the hypodermal syncytium without dividing; this alternate fate is represented by [S]). P(5-7).p divide to generate vulval cells. The lineage of each of these cells is represented by a set of four letters. Each letter indicates the fate of one of the granddaughters of P(5-7).p, referred to as a P(5-7).pxx cell. (An 'x' after a cell name refers to both daughters of that cell, for example the P5.px cells are P5.pa and P5.pp, the anterior daughter and the posterior daughter, respectively, of P5.p.) P6.p expresses a 1° lineage, which is represented by [TTTT] because the four P6.pxx cells (formed by two rounds of anterior-posterior divisions) all divide transversely (left, 1 and right, r). P5.p and P7.p, which express 2° lineages, undergo polar lineages, represented by [LLTN] and [NTLL], respectively; the lineage of P5.p is mirror-symmetric to that of P7.p. After P5.p undergoes two rounds of anterior-posterior divisions, each P5.pax cell divides longitudinally, [L]; P5.ppa divides transversely, [T]; and P5.ppp fails to divide, [N]. The lineage of P7.p is analogous. During the late L3 and early L4 stages, the anchor-cell-proximal cells, the descendants of P5.pp, P6.p and P7.pa (all cells formed by [T] or [N] lineages) detach from the ventral cuticle and move dorsally forming a vulval cavity (see ref. 8). The anchor-cell-distal cells, the descendants of P5.pa and P7.pp (all cells formed by [L] lineages) adhere to the ventral cuticle. (The character of adherence is indicated by the bold-face type of the 'L'). The anchor cell nucleus subsequently disappears and, at the L4 molt, certain of the invaginated cells (all daughters of P5.pp and P7.pa) move first toward the centre of the vulval cavity and then ventrally.



Here we report the vulval cell lineage defects resulting from representative alleles of genes defined by Muv and Vul mutants and assign 23 of these genes to specific steps in our recently-proposed pathway⁸ of wild-type vulval development.

Wild-type vulval development

A model for the pathway of development of the vulva of the wild-type *C. elegans* hermaphrodite is shown in Fig. 1^{1,5,7,8}. The vulva is formed by the descendants of three of six equipotent cells of the ventral hypodermis, in response to an inductive signal from the anchor cell of the somatic gonad. Vulval development consists of three temporally distinct steps (described in more detail below): the generation of the seven cells involved in vulval development, the determination of the fates of the equipotent vulval precursor cells in response to the anchor cell signal, and lastly the expression of each of these fates.

(1) **Generation.** The six cells with the potential to express vulval cell lineages, called P(3-8).p, are the posterior daughters of six of the twelve P cells. The nuclei of the twelve P cells migrate from a ventro-lateral position into the ventral cord during the middle of the first larval (L1) stage. Subsequently, each P cell divides to produce an anterior daughter (a Pn.a neuroblast) and a posterior daughter (a Pn.p hypodermal cell) (see Fig. 1a, which also explains the notation). Five of the Pn.p cells, P(1, 2, 9-11).p, fuse with the hypodermal syncytium during the L1 stage, but P(3-8).p, the 'vulval precursor cells', do not fuse. The gonadal anchor cell is formed during the second larval (L2) stage by either of two bipotent cells, Z1.ppp or Z4.aaa (Fig. 1a); interactions between these two cells cause one to become the anchor cell (ac) and the other to become a ventral uterine progenitor (VU) cell.

(2) **Determination.** Each of the six vulval precursor cells P(3-8).p is able to express any of three lineages (called 1°, 2° or 3°), each of which is defined by a distinct pattern of cell divisions (see below). During the early to middle third larval

(L3) stage a graded signal from the gonadal anchor cell induces vulva formation and specifies which of the three lineages each of the vulval precursor cells will express (Fig. 1b). The three cells P(5-7).p, which are near the anchor cell, are induced to generate the vulva: P6.p is determined to express a 1° lineage, whereas P5.p and P7.p are determined to express 2° lineages. The three cells P(3,4,8).p, which are farther from the anchor cell, are not induced and are determined to express a 3° lineage.

(3) **Expression.** During the middle to late L3 stage, the P(3-8).p cells express their lineages (1°, 2° or 3°); each lineage consists of a distinct pattern of cell divisions generating a particular set of cell types (Fig. 1c). The vulva is formed during the fourth larval (L4) stage by the descendants of P(5-7).p, which have expressed the 1° and 2° lineages.

Strategy for genetic analysis

To define a genetic pathway for the vulval cell lineages, we first observed the vulval cell lineages of individual hermaphrodites carrying an allele of each gene defined by a Muv or a Vul mutation. Description of the mutant phenotypes at the resolution of single cells allowed us to identify the first step in the developmental pathway visibly affected by a mutation in each gene. We confirmed these assignments by examining the patterns of epistasis among mutations postulated to affect different steps in the pathway. (One mutation is epistatic to a second mutation if the double mutant containing both mutations has the phenotype caused by the first mutation.) Because each step in the pathway of vulval development is dependent on the completion of the preceding step, a mutation that prevents the completion of an early step in the pathway should be epistatic to all mutations that prevent the completion of later steps. In addition, we determined the temperature-sensitive periods (TSPs) of those genes for which we had temperature-sensitive (ts) alleles. The TSP of a gene must be before, or concurrent with, all steps in the pathway for which that gene is needed.

Table 1 Lineages of P(3-8).p in mutants defective in the determination of precursor cell fates

Genotype	No. of animals	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p
Wild type		3°/S S S*	3° S S	2° LLTN	1° TTTT	2° NTLL	3° S S
<i>lin-10</i>		3°/S S S*	3° S S	3° S S	3°/1° S S	3° S S	3° S S
	8	S S*	S S	S S	S S	S S	S S
	3	S S*	S S	S S	TTTT	S S	S S
	2	S S*	S S	LLTN	TTTT	NTLL	S S
	1	S S	S S	LLTN	S S	NTLL	S S
	1	S S	S S	S TN	S S	NT S	S S
<i>lin-8; lin-9</i>		2°/1° LLOO	1°/2° TTTO	2°/1° LLLN	1°/2° TTTO	2° NOLL	2°/1° OTOO
	1	LLOO	TTTO	LLLN	TTTO	NOLL	OTOO
	1	LOOT	LTTO	LLTN	TTTT	NLLL	LOTO
	1	LLOO	OOLL	LLLN	TTTT	NLLL	LLTO
	1	TTOO	LLLN	TTTT	NLLL	LLOO	LLO?
<i>lin-13</i>		3°/2° LLOO	2°/u LLOO	2° LLON	1° T??T	2° NOLL	3°/u S S
	1	LLOO	LLOO	LLON	T??T	NOLL	S S
	1	S S	S S	LLTN	TTTT	NLLL	S OO
	1	S	OODL	LLTN	O??T	NTLL	S S
	1	OLOT	S OO	LLTN	TTTT	NTLL	S S
	1	S S	S LL	LLON	TTTT	NTLL	TTLS

The lineages of the vulval precursor cells P(3-8).p in Vul hermaphrodites of genotype *lin-10(e1439)*, in Muv hermaphrodites of genotype *lin-8(n111)*; *lin-9(n112)*, and in Muv hermaphrodites of genotype *lin-13(n387)*. In these strains, most of the aberrant lineages expressed by P(3-8).p were similar to one of the three lineages (1°, 2° or 3°) normally expressed by these cells in the wild type. Consequently, we have interpreted these lineage defects as transformations in the fates of the P(3-8).p cells. The top line for each gene shows a 'consensus fate' for each of the six cells; if a particular P(3-8).p cell expressed more than one fate, all fates observed in at least 25% of the animals examined are listed in the order of their frequencies and are separated by /. The subsequent lines for each gene describe the lineages of P(3-8).p in individual animals.

Lineage nomenclature is as follows. The lineage of each P(3-8).p cell is represented by one to four letters, each of which (with the exception of 'S'; see below) represents the fate of one of the four granddaughters of the P(3-8).p cell⁸. (Refer to Fig. 1 legend for the notation specifying cells.) The fate of each P(3-8).pxx cell comprises two parts: the axis of its division and the morphogenetic behaviours of its daughter cells. The axis of division is represented by individual letters (see Fig. 1 legend and below). The morphogenetic behaviour of the daughter cells is indicated by the style of the letter: if the daughters adhered to the ventral cuticle at the L3 molt, the letter is in boldface; if not, the letter is in roman typeface. We interpret the lineage of a P(3-8).p cell as a 1° lineage if all four P(3-8).pxx cells divided, represented by any of the letters 'T', 'O', 'L' or 'D' (see below), and if none of the eight descendant cells adhered to the ventral cuticle at the L3 molt, represented by roman typeface. We interpret the lineage of a P(3-8).p cell as a 2° lineage if at least three of the four P(3-8).pxx cells divided ('T', 'O', 'L' or 'D'), and if the four descendants of either P(3-8).pa or P(3-8).pb adhered to the ventral cuticle at the L3 molt, represented by two contiguous bold-face letters. We interpret the lineage of a P(3-8).p cell as a 3° lineage if neither of the P(3-8).px cells divided (a non-divided P(3-8).px cell is represented by 'S'). '1°', '2°', '3°', lineage with characteristics of the corresponding lineage of the wild type; 'u', uninterpretable lineage based on these criteria. Letters mean the following: 'N', 'T', 'L', as in the Fig. 1 legend; 'S', a P(3-8).p, P(3-8).px, or P(3-8).pxx cell appeared to join the hypodermal syncytium; 'O', a P(3-8).pxx cell divided with an oblique axis of division; 'D', a P(3-8).pxx cell divided, but its axis of division was not observed. '?', The fate of the P(3-8).pxx cell was not observed. Bold face indicates that the daughters of the P(3-8).pxx cell adhered to the ventral cuticle; otherwise, for 'T', 'L' and 'O', the daughters of the P(3-8).pxx cell moved dorsally within the hypodermis, and for 'N', the P(3-8).pxx cell moved dorsally within the hypodermis.

* In some animals P3.p did not divide and appeared to join the hypodermal syncytium.

Homoeotic defects in Muv and Vul mutants

We observed two general classes of vulval cell lineage defects in the Muv and Vul mutant strains. In strains of the first class, most of the mutant lineages expressed by the vulval precursor cells P(3-8).p are similar to one of the three wild-type lineages (1°, 2° or 3°) with regard to such criteria as time of cell divisions, axes of cell divisions and morphogenetic behaviour of descendant cells. We interpret such mutant lineages as transformations in the fates of the P(3-8).p cells. In Table 1, we present the vulval cell lineages of three such strains and describe the criteria used to classify the lineage of a particular P(3-8).p cell as a 1°,

2° or 3° lineage. In strains of the second class, most of the mutant lineages expressed by P(3-8).p are different from the normal 1°, 2° and 3° lineages. Nonetheless, many mutations of this second class also appear to cause transformations in the fates of the affected cells (see below).

In Table 2, we summarize the vulval cell lineages of Muv and Vul mutants defective in 22 genes. Most of these Muv and Vul mutations reduce or eliminate gene activity (refs 9, 11 and 12, summarized in Table 2), so the wild-type functions of most of these genes can be inferred from their mutant phenotypes.

Vulval cell lineages of mutants are variable

Unlike the vulval cell lineages of the wild type, those of the mutant strains we have studied are variable. The variability of the vulval cell lineages of some strains may result from residual gene activity (see Table 2), but such variability also occurs in strains in which we believe gene function has been eliminated. We have observed two types of variability. Some mutants have incompletely penetrant defects; for example in *lin-10* animals, P6.p expressed either the mutant fate (a 3° lineage) or the wild-type fate (a 1° lineage). We consider mutants of this class (mostly Vul mutants) to define genes that are involved in, and may be necessary for, the production of a particular cell fate or set of cell fates. In other mutants, a particular Pn.p cell can express one of several mutant fates; for example in *lin-8; lin-9* animals, P4.p, which expresses a 3° lineage in the wild type, expressed either a 1° lineage or a 2° lineage. (The Muv phenotype of *lin-8; lin-9* animals depends on both the *lin-8* and the *lin-9* mutations; see below.) We consider mutants of this class (mostly Muv mutants) to define genes that constrain a cell or set of cells to adopt a particular fate, so that if any such gene does not function, a cell may express any of a number of mutant fates.

We describe below how we have assigned 21 of the genes defined by Muv and Vul mutants, as well as two other genes involved in vulval development, to specific steps in the pathway of vulval development.

Genes for precursor cell generation

Mutations in five genes alter the fates of P(3-8).p before they divide in the L3 stage, indicating that these genes are involved in the generation of the vulval precursor cells. These five genes act at four temporally distinct steps. (1) Migration of P cell nuclei: mutations in *unc-83* and *unc-84* block the migrations of the nuclei of the P cells (the parents of P(3-8).p) into the ventral cord, preventing the subsequent development of the vulva^{10,13}. (2) Generation of hypodermal Pn.p cells: the *lin-26* mutation causes most of the 12 Pn.p cells to have a neural nuclear morphology during the L1 stage and other Pn.p cells to divide during the L1 stage generating two, three, or less frequently four neural-like progeny (H. Ellis, personal communication), transforming the fate of the Pn.p cells to a fate possibly similar to that of their sister cells, the Pn.a neuroblasts. (3) Generation of hypodermal cells with the potential to become vulval precursor cells: in animals homozygous for the mutation *n300*, the cells P(3-8).p generally do not divide and appear to fuse with the hypodermal syncytium, adopting fates characteristic of their homologues P(1, 2, 9-11).p. (4) Expression of L3-specific cell lineages: in *lin-4* animals, the P(3-8).p lineages are highly variable with respect to both the timing of divisions and the number of descendant cells¹⁴. The descendants of these lineages do not have morphologies characteristic of vulval cells. Based on its effect on other cell lineages, the *lin-4* mutation seems to cause P(3-8).p, which normally would express the L3-specific vulval fates, to express instead a fate (generation of neural and hypodermal cells) characteristic of their parent P cells during the L1 stage¹⁵.

To confirm that the genes involved in the generation of the vulval precursor cells act before genes involved in later steps of vulval development, we have examined the phenotypes of 33

Table 2 Genes that affect the vulval cell lineages

Gene	Allele(s) used	Phenotype	Nature of phenotype; effect on gene activity	Cell fates						P(1, 2, 9-11).p	No.
				P3.p	P4.p	P5.p	P6.p	P7.p	P8.p		
Wild type											
Generation mutants											
<i>lin-26</i>	<i>n156</i>	Vul	Recessive; reduction	n	n	n	n	n	n	n	9
—*	<i>n300</i>	Vul	Recessive; reduction or elimination	S	S	S	S	S	S	S	15
<i>lin-4</i>	<i>e912</i>	Vul	Recessive; reduction or elimination	S/r	r	r	r	r	r	S	3
Determination mutants											
<i>lin-2</i>	<i>e1309</i>	Vul	Recessive; probably elimination	3°/S	3°	3°	3°/u	3°	3°	S	9
	<i>e1453</i>										
<i>lin-3</i>	<i>e1417</i>	Vul	Recessive; reduction	3°/S	3°	3°	3°/1°	3°	3°	S	12
<i>lin-7</i>	<i>e1413</i>	Vul	Recessive; elimination	3°/S	3°	3°/2°/u	3°/1°	3°/u	3°	S	7
<i>lin-10</i>	<i>e1439</i>	Vul	Recessive; elimination	3°/S	3°	3°	3°/1°	3°	3°	S	15
<i>let-23</i>	<i>n1045</i>	Vul	Recessive; reduction	3°/S	3°	3°	3°	3°	3°	S	41
<i>lin-1</i>	<i>e1777</i>	Muv	Recessive; elimination	1°/u	2°	1°/2°	1°	2°	u/1°	S	6
	<i>e1275</i>										
<i>lin-17</i>	<i>n671</i>	Muv	Recessive; reduction or elimination	3°/S	3°	2°	1°	ab2°/2°	3°	S	9
	<i>n677</i>										
<i>lin-18</i>	<i>e620</i>	Muv	Recessive; elimination	3°/S	3°	2°	1°	ab2°/2°	3°	S	13
<i>lin-15</i>	<i>n309</i>	Muv	Recessive; reduction	2°	2°/1°	2°	1°	2°	2°/1°	S	5
<i>lin-13</i>	<i>n387</i>	Muv	Recessive; reduction	3°/2°	2°/u	2°	1°	2°	3°/u	S	5
<i>lin-8(n111);</i> <i>lin-9(n112)</i>		Muv	Recessive; reduction or elimination†	2°/1°	1°/2°	2°/1°	1°/2°	2°	2°/1°	S	4
<i>lin-12(0)</i>	<i>n137/</i> <i>n720</i>	Abn	Recessive; elimination	S/3°	u/3°	1°	1°	1°/u	3°	S	11
<i>lin-12(d)</i>	<i>n137</i>	Muv	Dominant; increased or ectopic expression	2°	2°	2°	2°	2°	2°	S	11
Expression mutants											
<i>lin-11</i>	<i>n382</i> <i>n389</i>	Vul	Recessive; reduction or elimination	S/3°	3°	ab2°	1°	ab2°	3°	S	7
Other mutants											
<i>lin-24</i>	<i>n432</i>	Vul	Dominant; novel function?	X/3°	X/3°	X/3°	X/1°	X/3°	X/3°	X/S	22
<i>lin-33</i>	<i>n1043</i>	Vul	Dominant; novel function?	X/3°	X/3°	X/3°	X/1°	X/3°	X/3°	X/S	15
<i>lin-34</i>	<i>n1046</i>	Muv	Dominant; unknown	(1°-2°)/3°	(1°-2°)	2°	1°	2°	(1°-2°)/3°	S	12
<i>lin-31</i>	<i>n301</i>	Muv	Recessive; elimination	d, (u/2°)	d, (u/2°)	d, (2°/u)	d, (1°/u)	d, (2°/1°)	d, (2°/3°)	S	11, 4
<i>lin-25</i>	<i>e1446</i>	Vul	Recessive; reduction or elimination	d, (3°)	d, (3°)	d, (3°)	d, (1°/u)	d, (3°)	d, (3°)	d/S	14, 5

The alleles listed are those used for the determination of the vulval cell lineages and/or for studies of gene interactions. The effects of Muv and Vul mutations on the activities of their respective genes are based on genetic analyses^{9,11,12}. The fates of P(1-11).p in animals of all genotypes except those listed below were determined using Nomarski optics to observe directly the divisions and patterns of differentiation of each of these cells and their descendants. The fates of P(1-11).p in *lin-34* animals, *let-23* animals, *lin-33* animals, some *lin-31* animals, and some *lin-24* animals were inferred from observations using Nomarski optics after the vulval cell divisions were complete. Animals of genotype *let-23* were grown at 15 °C before observation. The vulval cell lineages of *lin-4* animals are from ref. 14. For most strains, we examined the vulval anatomy of additional animals to confirm that the above lineages were representative of the defects in the strain. For each genotype, the consensus fates of P(1-11).p are described as in Table 1. In those strains, *lin-25* and *lin-31*, in which some of the cells P(1-11).p divided before the L3 stage, indicated by 'd' the fates shown in parentheses are those the P(1-11).p cell or its descendants adopted during the L3 stage. 'Generation' mutants define genes necessary for the generation of P(3-8).p cells with the potential to express vulval cell lineages (see text). 'Determination' mutants define genes necessary for the proper determination of the fates of the P(3-8).p cells in response to the inductive signal from the anchor cell. In such mutants, some of the P(3-8).p cells express 1°, 2°, or 3° lineages that are normally expressed only by others of these cells. The criteria used to distinguish 1°, 2°, and 3° lineages are described in the legend to Table 1. For general descriptions of the vulval cell lineages of animals of genotypes *lin-1*, *lin-2*, *lin-3*, *lin-7* and *lin-8*; *lin-9*, see ref. 10. The vulval cell lineages of animals of genotypes *lin-12(d)*, *lin-12(0)*, *lin-15*, *let-23* and additional animals of genotypes *lin-2*, *lin-3* and *lin-7* will be presented elsewhere (P.W.S. and H.R.H., in preparation). 'Expression' mutants define genes necessary for the expression of one of the three vulval fates (see text). We were unable to assign 'Other' mutants to specific steps in the pathway of vulval development (see text). For *lin-25* and *lin-31*, the first number is the number of animals observed during the L1 or L2 stage to confirm the presence of Pn.p divisions before the L3 stage, and the second number is the number of animals for which vulval cell lineages were observed during the L3 stage. 'Muv', multivulva; 'Vul', vulvaless; 'Abn', abnormal vulva. Descriptions: '1°', '2°', '3°', 'u', as in Table 1; other entries as follows: '(1°-2°)', by anatomical criteria expressed a 1°-like or 2°-like lineage, but the lineage was not directly observed; 'ab2°', underwent an abnormal 2° lineage (see text); 'S', did not divide and appeared to join the hypodermal syncytium; 'd', divided precociously; 'n', the nucleus had a compact neural appearance or generated progeny with compact neural nuclei; 'X', had abnormal morphology in the L1 stage and sometimes died; 'r', divided variably, possible reiteration of ancestral fate (see text).

* —Here *n300* refers to the mutation associated with the Vul phenotype observed in animals homozygous for the reciprocal translocation *nT1(IV;V)*¹². Because no single gene mutation has been isolated that fails to complement the Vul phenotype caused by *n300*, no gene name has been given to this mutation.

† —The Muv phenotype of *lin-8*; *lin-9* animals is dependent on the presence of both the *lin-8* and the *lin-9* mutations, each of which alone confers a wild-type phenotype (ref. 9; also, we examined at the level of single cells the vulval anatomy of ~20 animals of each of the two genotypes; all *lin-8* and all *lin-9* animals were wild-type).

Table 3 Gene interactions

	Determination Muv				
	<i>lin-1</i>	<i>lin-12</i> (<i>n137</i>)	<i>lin-15</i>	<i>lin-8</i> ; <i>lin-9</i>	<i>lin-13</i>
Generation Vul					
<i>lin-26</i>					
<i>n300</i>	Vul*	Vul	Vul	Vul	Vul
<i>lin-4</i>					
<i>lin-24(n432)</i>					
Determination Vul					
<i>let-23</i>	Muv	Muv	Muv/Vul	Vul	Vul
<i>lin-2</i>					
<i>lin-7</i>	Muv	Muv	Muv	Muv/Vul	Vul
<i>lin-10</i>					
<i>lin-3</i>	Muv	Muv	Muv	Muv	Muv

The interaction between a multivulva and a vulvaless mutation was examined using a dissecting microscope to determine the vulval anatomy of a strain that contained both mutations and to compare the gross phenotype of the double mutant to that of each single mutant, which has been described¹². Double mutant strains were constructed by picking ~10 homozygous Muv and 10 homozygous Vul animals from the progeny of doubly heterozygous hermaphrodites. As some of the Muv or Vul animals that were picked were also heterozygous for a mutation of the opposite phenotype (for example, some homozygous Muv animals were heterozygous for the Vul mutation), the double mutant strain was obtained in the next generation by picking a segregant of a Muv or Vul animal that had a phenotype different from that of its parent (such as a Vul animal derived from a Muv parent). The Muv and Vul alleles used are listed in Table 2: *lin-1(e1275)* was used to construct most double mutants containing *lin-1*, and *lin-2(e1309)* was used to construct most double mutants containing *lin-2*. In a few double mutants, alternative alleles were used; *lin-2(e1453)* was used for the *lin-12(d)*; *lin-2* and *lin-2 lin-15* strains; *lin-10(e1438)* was used for the *lin-10*; *lin-8*; *lin-9* strain; and *lin-1(e1777)* was used for the *let-23*; *lin-1* strain. Other than those listed, additional double mutants containing a mutation that affected the generation of the vulval precursor cells and a mutation that affected a later step in the pathway of vulval development were also constructed. In all cases the mutation that affected the generation of the vulval precursor cells was epistatic to the mutation that affected the later step. Specifically, *unc-84(e1410)* was epistatic to mutations in genes affecting the determination of the fates of the precursor cells—*let-23*, *lin-12(d)*, and *lin-8*; *lin-9*—as well as to mutations in genes affecting the expression of the vulval lineages—*lin-11* and *lin-18*. In addition, the vulval cell lineage defects resulting from *n300* were present in double mutants carrying *n300* and Vul mutations in any of three genes—*lin-2*, *lin-7*, and *lin-10*—that affect the determination of the fates of the precursor cells, and the vulval cell lineage defects resulting from *lin-24* were present in *lin-24*; *lin-2* and *lin-10*; *lin-24* double mutants. Lastly, *lin-26*, *n300*, *lin-4*, and *lin-24* were epistatic to the Muv mutation *lin-31*. Muv, the double mutant strain had a Muv phenotype; Vul, the double mutant strain had a Vul phenotype; Muv/Vul, the phenotypes resulting from the Muv and Vul mutations were co-expressed in the double mutant strain. In Muv/Vul strains, some animals had Muv characteristics, some were Vul, and some animals had both Muv and Vul characteristics. We examined animals of some such strains using Nomarski optics; in certain animals of each strain examined, some of the cells P(3,4,8).p, which express a 3° lineage in the wild type, expressed 1° or 2° lineages, while some of the cells P(5-7).p, which express a 1° or 2° lineage in the wild type, expressed 3° lineages.

* As the translocation *nT1(IV;V)* suppresses recombination in an interval of LGIV that includes *lin-1*¹², the double mutant containing *lin-1* and the mutation causing the Vul phenotype associated with *nT1*, *n300*, was not constructed.

strains carrying a mutation affecting the generation of the vulval precursor cells and a mutation postulated to affect a later step. In all of these strains, the mutation affecting the generation of the vulval precursor cells was epistatic to the second mutation in the strain (Table 3). For example, the Vul mutation *n300*, which causes the vulval precursor cells to join the hypodermal syncytium without dividing, is epistatic to all Muv mutations that cause the six vulval precursor cells to express 1° and 2°

vulval lineages. Thus, *n300* disrupts a step necessary for the generation of cells capable of expressing vulval lineages.

A gene controlling anchor cell generation

Greenwald *et al.*¹¹ have shown that the gene *lin-12* controls the generation of the anchor cell. Recessive *lin-12(0)* loss-of-function mutations cause both potential anchor cells (Z1.ppp and Z4.aaa in Fig. 1a) to become anchor cells. Conversely, dominant *lin-12(d)* elevation-of-function mutations cause both of these cells to become VU cells. Thus, the level of *lin-12* activity specifies whether a potential anchor cell will become an anchor cell or a VU cell. (Abnormalities in anchor cell number caused by *lin-12* mutations have an effect on the fates of the vulval precursor cells distinct from the direct action of *lin-12* in specifying the fates of these cells; see below and P.W.S. and H.R.H., in preparation.)

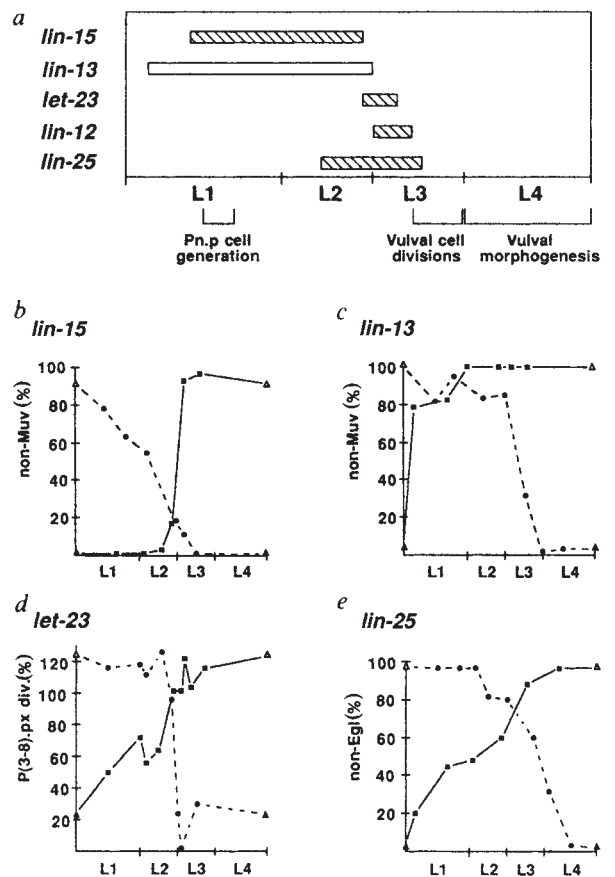
Genes for precursor cell determination

We have identified 15 genes that affect the determination of which of the three alternative vulval lineages each precursor cell expresses. Reduction-of-function mutations in any of these genes cause some precursor cells to adopt fates characteristic of other precursor cells. Three classes of transformations in cell fates are known (Table 1, ref. 10; P.W.S. and H.R.H., in preparation; summarized in Table 2). (1) In Vul mutants defective in *lin-2*, *lin-3*, *lin-7*, *lin-10* or *let-23*, the P(5-7).p cells generally express 3° lineages. Thus, these genes are necessary for the generation of the 1° and 2° lineages (see Fig. 1c for a description of each lineage). (2) In Muv mutants defective in *lin-1*, *lin-15* or *lin-13*, P(3,4,8).p often express 1° or 2° lineages. Animals of genotype *lin-8*; *lin-9* have a similar Muv phenotype. However, the Muv defect of this strain is dependent upon two mutations: either the *lin-8* or the *lin-9* mutation alone confers a wild-type phenotype (ref. 9 and Table 2 legend). In addition, mutations in any of four other genes—*lin-35*, *lin-36*, *lin-37* and *lin-38*—result in a wild-type phenotype on their own but generate a Muv phenotype with a *lin-8* or a *lin-9* mutation (E.L.F. and H.R.H., in preparation). Thus, we have identified nine genes with activities that, either singly or in certain pairwise combinations, appear necessary for the generation of the 3° lineage. These genes may act by promoting the 3° lineage and/or by repressing the 1° and 2° lineages. (3) Recessive loss-of-function *lin-12(0)* mutations cause P(3-8).p generally to express 1° or 3° lineages, but never 2° lineages¹¹. Conversely, dominant elevation-of-function *lin-12(d)* mutations, which result in a Muv phenotype, cause P(3-8).p to express 2° lineages¹¹. Thus, the level of *lin-12* activity specifies whether a P(3-8).p cell will express a 2° lineage or a non-2° lineage.

Site of action of precursor cell determination genes. The determination of the fates of the vulval precursor cells involves the production of an inductive signal by the gonadal anchor cell and the response to it by the hypodermal vulval precursor cells. Thus, mutants abnormal in the determination of precursor cell fates could be defective either in the gonad or in the hypodermis. For example, the expression of 1° and 2° lineages by P(3,4,8).p in the Muv mutants could be caused by (1) a defect in the gonad resulting in an increase in the strength of the inductive signal, (2) a defect in the hypodermis resulting in an increased sensitivity to the inductive signal, or (3) a defect in the hypodermis resulting in the expression of 1° and 2° lineages independently of an inductive signal. To determine whether any of the Muv mutations are of the third class, we have used a laser microbeam^{5,16} to ablate the gonads of five L1 animals of each of the Muv genotypes *lin-1*, *lin-12(d)*, *lin-13*, *lin-15*, and *lin-8*; *lin-9*. (The ablation of the gonad of an L1 wild-type animal causes all six vulval precursor cells to assume a 3° fate⁵.) All 25 mutant animals became Muv, indicating that the expression of 1° and 2° lineages in these strains is independent of the gonadal inductive signal. Thus, the site of action of these signal-independent Muv mutations seems likely to be within the hypodermal vulval

Fig. 2 Temperature-shift experiments. *a*, Temperature-sensitive periods (TSPs) (hatched rectangles) for *lin-15*(n765), *let-23*(n1045), *lin-12*(n137 n460), and *lin-25*(n545) activities; these TSPs are the minimal periods based on our data during which the gene product must be active to generate a wild-type phenotype. The open rectangle indicates the period, which is not strictly a TSP (see below), during which *lin-13*(n387) animals must be at the restrictive temperature to display an essentially mutant phenotype. The beginning of the TSP is defined²⁴ as the earliest point at which a shift from the restrictive to the permissive temperature results in a higher percentage of mutant animals than that observed in animals maintained at the permissive temperature. Similarly, the end of the TSP is defined²⁴ as the latest point at which a shift from the permissive temperature to the restrictive temperature results in a higher percentage of mutant animals than that observed in animals maintained at the permissive temperature. The TSP of each gene could encompass a slightly larger period of time than shown in the diagram, in that the beginning of each TSP could extend up to, but not include, one point earlier than indicated and the end could extend up to, but not include, one point later than indicated. We have chosen not to present this larger interval because we have no data that indicate that these genes act beyond the periods shown. The TSP of *lin-12*(n137 n460) is based on ref. 11. Panels *b-e*: (■ and solid lines), percentage of animals shifted from the permissive to the restrictive temperature at the stage indicated that expressed a non-mutant phenotype; (● and dotted lines), percentage of animals shifted from the restrictive to the permissive temperature at the stage indicated that expressed a non-mutant phenotype. Δ, Percentage of animals grown at the permissive temperature that expressed a non-mutant phenotype. ▲, For *lin-13*, *lin-15* and *lin-25*, the percentage of animals shifted from the permissive to the restrictive temperature before hatching that expressed a non-mutant phenotype; for *let-23*, the percentage of animals grown at the restrictive temperature that expressed a non-mutant phenotype.

Methods. In *b*, stocks of *lin-15*(n765) animals were grown at the permissive temperature of 15°C. A synchronous population was obtained by placing three or four adult hermaphrodites on each of 36 Petri dishes equilibrated at 15°C and allowing these animals to lay eggs for two hours. Immediately after removal of the parents, half the Petri dishes were shifted to the restrictive temperature of 25°C and half were kept at 15°C. At various later times, two Petri dishes, containing between 23 and 86 animals, were shifted either from 15°C to 25°C or from 25°C to 15°C. In most cases, the stage of development of three or four of the animals that were shifted was ascertained using Nomarski optics¹; otherwise, the stage of the animals at the time of transfer was inferred from animals observed at other time points during the same experiment. Two Petri dishes at each temperature were not shifted. 'non-Muv (%)', percentage of *lin-15* animals that did not have multiple pseudovulvae after reaching adulthood. (The wild type is 100% non-Muv.) In *c*, because *lin-13*(n387) results in a sterile Muv phenotype, hermaphrodites of genotype *+lin-13 unc-36(e251)/+lin-16(e1743)+unc-86(e1507)* were grown at the permissive temperature of 15°C. The temperature-shift experiments were performed on the progeny of these animals as in 'b'. The temperature-shift data indicate that most (>75%) *lin-13* animals have essentially wild-type phenotypes if they are grown at the permissive temperature for any period of time between the early L1 stage and the L2 molt. This 'temperature-sufficient period' of *lin-13* could result from either of two causes²⁵. First, the *lin-13* mutation could be temperature-sensitive for the synthesis of its gene product, with the product made in excess continuously from the early L1 until the L2 molt, after which it functions. Alternatively, *lin-13* could have two TSPs, one ending at the beginning of the temperature-sufficient period (the early L1) and the other starting after that time and ending at the end of the temperature-sufficient period (the L2 molt), with *lin-13(+)* activity during either of the TSPs sufficient to generate a wild-type phenotype. The latter hypothesis is consistent with the observation¹² that at the permissive temperature *lin-13(+)* acts both maternally and zygotically. 'non-Muv (%)', as in *b*. In *d*, stocks of *let-23*(n1045) animals were grown at the permissive temperature of 25°C and at the restrictive temperature of 15°C. Individual animals were shifted from 25°C to 15°C or from 15°C to 25°C. The stage of development of each animal was ascertained using Nomarski optics at the time of the shift^{1,6}. Between 1 and 15 animals were shifted at each time point. 'P(3-8).px div. (%)', the number of P(3-8).px cells that divided in individual animals compared to the six (P(5-7).px) that divide in the wild type. (The wild type is 100% P(3-8).px div.) At the permissive temperature, *let-23*(n1045) causes more P(3-8).px cells to divide than in the wild type; this phenotype, the basis of which will be discussed elsewhere, is likely to result from a small decrease in *let-23* gene activity (P.W.S. and E.L.F., unpublished observations.) In *e*, stocks of *lin-25*(n545) animals were grown at the permissive temperature of 15°C and temperature-shift experiments were performed as in *b*. 'non-Egl (%)', percentage of adult *lin-25* animals competent to lay eggs. (The wild type is 100% non-Egl.) As hermaphrodites with certain vulval cell lineage defects nonetheless can lay eggs (ref. 10 and P.W.S., unpublished observations), the egg-laying competent *lin-25* animals may not have expressed entirely wild-type vulval lineages.



precursor cells.

Similarly in Vul mutants, P(5-7).p could adopt 3° lineages either because the activity of the inductive signal of the gonad is reduced or because the ability of the precursor cells to respond to the inductive signal is reduced. By examining the phenotypes of double mutants carrying a Vul mutation and any of the signal-independent Muv mutations described above, we can distinguish certain classes of Vul mutations. Just as the physical ablation of the gonad does not affect the Muv phenotype of these mutants, the 'genetic ablation' of the gonadal signal by a gonad-specific Vul mutation should not affect the phenotype caused by a signal-independent Muv mutation. Conversely, if a Vul mutation has a site of action in the hypodermis, the phenotype of the Vul mutation could be epistatic to, co-expressed with (see Table 3 legend), or not alter the phenotypic effect of a signal-independent Muv mutation. Thus, if a Vul mutation affects the Muv phenotype caused by any signal-independent Muv mutation, the Vul mutation must have a site of action outside the gonad and thus presumably in the hypodermis. However, if all signal-independent Muv mutations are epistatic to a Vul mutation, the site of action of the Vul mutation could be either in the gonad or in the hypodermis.

Vul mutations in *let-23*, *lin-2*, *lin-7* and *lin-10* affected the Muv phenotypes of at least some hypodermal-acting Muv mutations (Table 3), indicating that mutations in these four genes have sites of action in the hypodermal precursor cells. The site of action of *lin-3* was not revealed by these experiments (all signal-independent Muv mutations were epistatic to the Vul allele *lin-3(e1417)*), but the Vul phenotype resulting from a further reduction in *lin-3* activity was co-expressed with the signal-independent Muv phenotype resulting from a less severe *lin-15* allele, indicating that *lin-3* also affects the hypodermis. Specifically, at 20°C the Muv phenotype of animals of genotype *lin-3(n378)/lin-3(n1059); lin-15(n765ts)* is less severe than the Muv phenotype of *lin-15(n765ts)* animals. (Note that *n378* is a Vul allele of *lin-3*, *n1059* is a putative null allele of *lin-3* and results in larval lethality, and *n765ts* is a *lin-15* allele that at 20°C results in a less expressive Muv phenotype than the reference allele of *lin-15*, *n309*; ref. 12.)

Thus, all mutations that affect the determination of the fates of the precursor cells appear to have a site of action in these cells, implying that these gene products are involved in the reception of and/or response to the inductive signal. We have failed to identify any genes in which loss-of-function mutations

Table 4 Lineages of P(5-7).p in mutants defective in expression of 2° lineages

Genotype	No. of animals	P5.p	P6.p	P7.p
Wild type		2° LLTN	1° TTTT	2° NTLL
<i>lin-11</i>		ab2°	1°	ab2°
	1	LLLL	TOOT	LLOO
	1	LLLO	TTOT	OLOL
	1	LLLL	OTTO	OOLL
	1	LLL?	TTTT	TLLO
	1	LLLO	TTOT	OOLL
	1	LOLO	TTOT	OLL
	1	OLLO	TTTT	LLOL
<i>lin-17</i>		2°	1°	ab2°
	4	LLTN	TTTT	LLLL
	1	LLTN	TTTT	LLTN
	1	LLTT	TTTT	LLLL
	1	LLLT	TTTT	LLTT
	1	LLTN	TTTT	TTLL
	1	LLTN	TTTT	LLOO
	1	LLTT	TTTT	LLLL
	1	LLTT	TTTT	LLLO
	1	LLTT	TTTT	TOLL
<i>lin-18</i>		2°	1°	ab2°/2°
	3	LLTN	TTTT	NTLL
	2	LLTN	TTTT	LLNT
	1	LLTN	TTTT	TTLL
	1	LLTN	TTTT	LLTN
	1	LLTN	TTTT	LLNO
	1	LLTN	TTTT	LLLL
	1	LLTN	TTTT	TOLL
	1	LLLL	TDTT	LLLL
	1	LLTO	TTTT	LLLL
	1	LLTT	TTTT	NTLL

The lineages of P(5-7).p in *lin-11*(n382, n389) hermaphrodites, *lin-17*(n671, n677) hermaphrodites and *lin-18*(e620) hermaphrodites are represented using the nomenclature of Table 1 legend. (In all animals of these genotypes, the other three precursor cells P(3,4,8).p underwent lineages identical to those expressed by the equivalent cells in the wild type.) In these strains, the lineages of P5.p and P7.p, the cells determined to express a 2° lineage in the wild type, can be abnormal. The top line of each entry summarizes the 'consensus fates' expressed by P(5-7).p in each strain (see the legend to Table 1). The subsequent lines of each entry contain the lineages of P(5-7).p in individual animals of each strain. For *lin-11*, the first four lines represent lineages of *lin-11*(n382) animals, and the remaining three lines represent lineages of *lin-11*(n389) animals. In *lin-11* animals, all four P5.p and P7.p daughter cells express an [LL]-like pattern of division. (Although some P5.pxx or P7.pxx cells in *lin-11* animals divided with an oblique axis, [O], their descendants adhered to the ventral cuticle, as do the descendants of an [L] lineage.) For *lin-17*, the first six lines represent lineages of *lin-17*(n671) animals, and the remaining three lines represent lineages of *lin-17*(n677) animals. Note that *lin-18*(e620) results in a slightly temperature-sensitive phenotype. For *lin-18*, the first five lines represent lineages of *lin-18* animals grown at 20°C. The remaining five lines represent lineages of *lin-18* animals that were grown at 25°C until the P(5-7).pxx cells had divided. These animals were then placed at 20°C and the final round of cell divisions was observed; cell lineages were inferred based upon the positions of the P(3-8).pxx cells in individual animals. In some *lin-17* and *lin-18* animals, the lineage of P7.p was symmetric but the fates expressed by all four P7.pxx cells were abnormal (the P7.pxx cells divided with a longitudinal axis, but their daughter cells did not adhere to the ventral cuticle, represented by [LLLL]). Descriptions: '1°', '2°', 'L', 'T', 'N', 'O', 'D', '?', as in Table 1 legend; 'ab2°', expressed abnormal 2° lineages with a defect characteristic of the particular mutant strain (see text).

eliminate the inductive signal of the anchor cell. Perhaps the genes that function in the production of the signal have essential functions and a loss-of-function mutation in any such gene results in lethality. Alternatively, it may not be possible to eliminate the inductive signal by a mutation in a single gene, either because the genes that control the signal are multigene family members or because the signal itself has multiple at least partially redundant components. It also remains possible that any of the genes classified as involved in the response to the signal could also affect its production. In our experiments, any mutation that caused a defect in both the production of and the response to the signal would be scored simply as defective in the response to the signal. Moreover, one or more of the genes defined by Vul mutations may normally act in the gonad but appear to be hypodermal-acting by our tests; in particular, a mutation might result in a Muv phenotype by causing such a gene to be active in the hypodermis, in which case the Vul mutation would be epistatic to the Muv mutation (because in the Muv mutant background the gene defined by the Vul mutation would indeed be acting in the hypodermis).

Times of action of precursor cell determination genes. We have performed temperature-shift experiments using temperature-sensitive alleles of three genes we believe to be involved in the determination of the precursor cell fates—heat-sensitive alleles of *lin-13* and *lin-15* and a cold-sensitive allele of *let-23*. The results are consistent with the hypothesis that these genes are involved in the determination of the fates of the precursor cells as opposed to the expression of those fates (Fig. 2). Specifically, the period during which the temperature-sensitive activity of each gene is needed to generate a wild-type phenotype ends before the onset of vulval cell divisions in the middle of the L3 stage. If the TSP of any of these genes had extended beyond the onset of vulval cell divisions (as is the case for *lin-25*; see below), we would have concluded that that gene acts after the period in which vulval precursor cell fates are determined.

Our temperature-shift data also suggest that these three genes, as well as *lin-12*¹¹, act during different periods of larval development. Specifically, *lin-15* and possibly *lin-13* appear to function from the L1 stage until the late L2 stage, *let-23* appears to function from the late L2 stage to the early L3 stage and *lin-12* appears to function during the early L3 stage. However, if the TSPs of any of these genes reflect temperature-sensitive defects in the synthesis of gene products (Fig. 2, legend), the results of these experiments simply indicate that these gene products are synthesized before the onset of vulval cell divisions and may function at any time after these divisions.

Genes for 2° lineage expression

We have identified three genes that are necessary for the expression of the 2° lineages. In the wild type, the two vulval precursor cells that are determined to undergo a 2° lineage, P5.p and P7.p, express asymmetric lineages of opposite polarity: P5.p undergoes a lineage represented as [LLTN], and P7.p undergoes a mirror-symmetric lineage represented as [NTLL]. (Each letter represents the fate of a P(3-8).pxx cell, one of the granddaughters of the P(3-8).p cell; see the legend to Fig. 1, and ref. 8 for details.) The first division of the 2° lineage is asymmetric, as the two cells generated express different fates, [LL] and [TN].

In mutants defective in *lin-11*, *lin-17* or *lin-18*, the asymmetry of the 2° lineages is disrupted (Table 4). In *lin-11* mutants, the 2° lineages are symmetric; all four P5.p and P7.p daughter cells express the [LL] portion of the 2° lineage, so that many P5.p and P7.p cells undergo [LLLL] lineages. In *lin-17* and *lin-18* mutants, the asymmetry of the P7.p lineage can be disrupted in two ways. In some *lin-17* and *lin-18* animals, the lineage expressed by P7.p is symmetric, but each P7.p cell expresses a novel fate [LL]. (The fate [LL] indicated with bold type is distinct from the fate [LL] indicated in non-bold type; see Table 4, legend.) In other animals of these genotypes, the polarity of the P7.p lineage is reversed, so that P7.pa rather than its sister

Table 5 Specificity for 2° lineages exhibited by expression mutants

	Wild type						<i>lin-11</i> , <i>lin-17</i> or <i>lin-18</i>					
	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p	P3.p	P4.p	P5.p	P6.p	P7.p	P8.p
Wild type	3°/S	3°	2°	1°	2°	3°	3°/S	3°	2°/ab2°	1°	ab2°/2°	3°
<i>lin-12(d)</i>	2°	2°	2°	2°	2°	2°	ab2°/2°	ab2°/2°	ab2°/2°	ab2°/2°	ab2°/2°	ab2°/2°
<i>lin-12(0)</i>	S/3°	u/3°	1°	1°	1°/u	3°	3°/S	3°/u	1°	1°	1°/u	3°
<i>ac</i> ⁻	3°/S	3°	3°	3°	3°	3°	3°/S	3°	3°	3°	3°	3°

The expression of the *lin-11*, *lin-17* and *lin-18* phenotypes in animals in which all, some, or none of the P(3–8).p cells are determined to express a 2° lineage. In *lin-12(d)* animals, the 2° lineage of each of the vulval precursor cells has a characteristic polarity (P.W.S. and H.R.H., in preparation): P(3–5).p express 2° lineages with the polarity of the wild-type P5.p lineage, P6.p expresses a 2° lineage with the polarity of the wild-type P7.p lineage, and P7.p and P8.p can express a 2° lineage of either polarity. In *lin-11*; *lin-12(d)* hermaphrodites (6 animals), all six precursor cells expressed [LLLL]-like lineages. In *lin-17*; *lin-12(d)* hermaphrodites (5 animals) and *lin-12(d)*; *lin-18* hermaphrodites (2 animals), all six precursor cells either expressed a 2° lineage with the polarity of P5.p or expressed a symmetrical but abnormal lineage [LLLL], which is characteristic of *lin-17* and *lin-18* animals (see text). In *lin-12(0)* animals, none of the six precursor cells is determined to express a 2° lineage. In *lin-11*; *lin-12(0)* hermaphrodites (10 animals), *lin-17*; *lin-12(0)* hermaphrodites (5 animals), or *lin-12(0)*; *lin-18* hermaphrodites (4 animals), none of the six precursor cells expressed any of the defects associated with *lin-11*, *lin-17* or *lin-18*. (The vulval cell lineages of *lin-11*; *lin-12(0)* animals were not determined directly; instead, these animals were found to be similar in vulval anatomy to *lin-12(0)* animals, indicating that the *lin-11* defect probably is not expressed in a *lin-12(0)* animal.) *lin-11*, *lin-17* and *lin-18* animals that lacked an anchor cell were obtained by ablating the gonads of seven *lin-11*, five *lin-17*, and seven *lin-18* L1 animals with a laser microbeam. Descriptions '1°', '2°', '3°', 'u', 'S', as in the legend to Table 1; 'ab2°', expressed an abnormal 2° lineage characteristic of *lin-11*, *lin-17*, or *lin-18*.

cell P7.pp undergoes an [LL] lineage. Note that *lin-17* mutations also cause at least five other blast cells that normally undergo asymmetric cell divisions to produce daughter cells with identical, although sometimes abnormal, fates (P.W.S. and H.R.H., in preparation). Thus, it is likely that mutations in *lin-11*, as well as mutations in *lin-17* and possibly *lin-18*, cause certain sister cells that normally adopt different fates to adopt similar fates instead.

If these three genes specifically affect the expression of the 2° lineages, the defects caused by mutations in these genes should be observed in all vulval precursor cells that are determined to express a 2° lineage and should not be observed in any vulval precursor cells determined to express a 1° or a 3° lineage. We tested this hypothesis by observing the expression of the *lin-11*, *lin-17* and *lin-18* defects in animals in which either all or none of the six precursor cells were determined to express 2° lineages (Table 5). First, a *lin-12(d)* mutation causes all P(3–8).p cells to be determined to express a 2° lineage (see above); in double mutants carrying a *lin-12(d)* mutation and a *lin-11*, *lin-17* or *lin-18* mutation, all six P(3–8).p cells displayed the defects

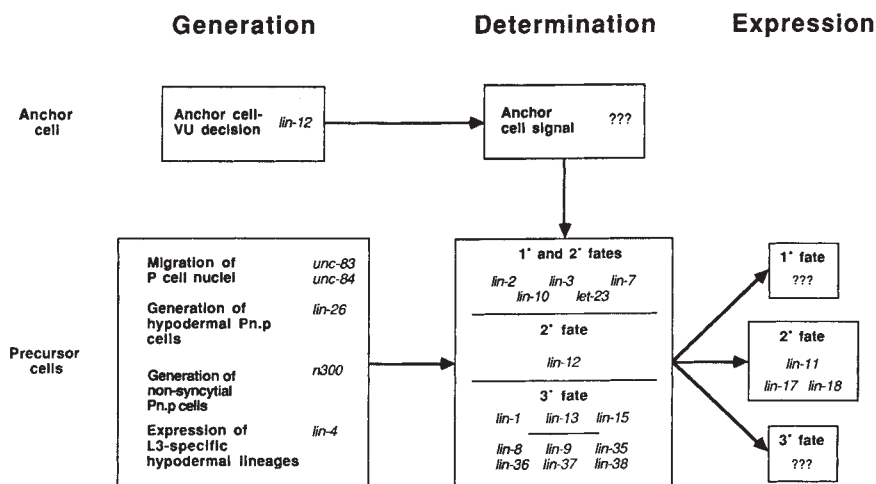
associated with a *lin-11*, *lin-17* or *lin-18* mutation. Second, a *lin-12(0)* mutation causes all P(3–8).p cells to be determined to express non-2° lineages; in double mutants carrying a *lin-12(0)* mutation and a *lin-11*, *lin-17* or *lin-18* mutation, none of the six P(3–8).p cells displayed any of the defects associated with a *lin-11*, *lin-17* or *lin-18* mutation. Third, ablation of the gonad in L1 wild-type animals causes all P(3–8).p cells to be determined to express 3° lineages; in gonad-ablated *lin-11*, *lin-17* and *lin-18* animals, all six P(3–8).p cells expressed a 3° fate.

Thus, *lin-11*, *lin-17* and *lin-18* appear to be required specifically for the expression of the 2° lineage. Furthermore, the determination of the fates of the vulval precursor cells and the expression of the vulval lineages (or at least the expression of the 2° lineage) are independent, genetically separable steps.

Genes not assigned to specific steps

We have been unable to assign five genes to specific steps in the pathway of vulval development. Three of these genes—*lin-24*, *lin-33* and *lin-34*—are defined by partially dominant mutations that cause defects in particular steps in the pathway. However,

Fig. 3 A genetic pathway for the specification of the vulval cell lineages (see text). The postulated order of the five genes involved in the generation of the vulval precursor cells is indicated from top to bottom within the box. This order is based only upon mutant phenotypes; epistasis experiments have not been performed. Mutations in two other genes besides *lin-4* cause the vulval precursor cells to adopt fates normally expressed at other times during development¹⁵. Dominant mutations in *lin-14* cause a Vul phenotype identical to that observed in *lin-4* animals; recessive mutations in *lin-14* or *lin-28* cause P(3–8).p to express vulval-like cell lineages during the L2 stage. However, as the activities of *lin-14* and *lin-28* are not necessary for the expression of vulval cell lineages, we have omitted these genes here. The nine genes necessary for the expression of the 3° lineage are divided into two classes. The genes *lin-1*, *lin-13* and *lin-15* have been identified by mutations that result in a Muv phenotype; *lin-8*, *lin-9*, *lin-35*, *lin-36*, *lin-37* and *lin-38* have been identified by mutations that alone result in a wild-type phenotype but in certain pairwise combinations result in a Muv phenotype (see text). The gene *lin-12* acts twice in the pathway. First, *lin-12* is involved in the specification of the fates of the two bipotent cells, Z1.ppp and Z4.aaa, one of which becomes the anchor cell. Second, *lin-12* activity is necessary for the vulval precursor cells to adopt 2° lineages.



there are no known loss-of-function alleles of these genes¹². Because dominant mutations can cause a gene product to be expressed inappropriately or to assume a novel function, we do not know if the wild-type products of these genes are involved in vulval development. In *lin-24* and *lin-33* animals, which have a Vul phenotype, up to six of the 12 Pn.p cells die during the late L1 or early L2 stage (Table 2). (The surviving P(3-8).p cells sometimes express abnormal cell lineages during the L3 stage, moving dorsally within the ventral cord and generating up to eight nuclei with morphologies characteristic of neural cells.) These results suggest that the *lin-24* and *lin-33* mutations result in products with cytotoxic activities that disrupt the generation of the vulval precursor cells. In *lin-34* animals, which have a phenotype similar to that of *lin-15* or *lin-8*; *lin-9* animals, P(3,4,8).p express 1° or 2° lineages, but P(5-7).p express their normal lineages (Table 2), indicating that this mutation disrupts the determination of the 3° fate.

Our observations concerning two other genes—*lin-25* and *lin-31*—are inconsistent with their involvement at only a single step in the pathway of vulval development. In *lin-25* animals, which have a Vul phenotype, some of the Pn.p cells divide once during the L1 stage. In the L3 stage, most of the vulval cells (P(3-8).p or their L1-derived daughters) adopt 3° fates. The L1 defect of this mutant indicates that *lin-25* is active during the L1 stage in the formation of the vulval precursor cells. However, the TSP of *lin-25* extends to the middle of the L3 stage (Fig. 2), past the time at which the P(3-8).p cells divide, indicating that *lin-25* must also be involved in the expression of the vulval cell lineages. In *lin-31* animals, which have a Muv phenotype, some of the P(3-8).p cells divide during the late L2 stage to produce two to four descendants. In the L3 stage, the P(3-8).p cells or their L2-derived progeny can undergo three additional rounds of division, generating either vulval-like cells or compact neural-like cells. Thus, *lin-31* also may act at more than one of the steps of vulval development.

Conclusions

We have determined that the defects in vulval morphology resulting from mutations in any of 28 genes are caused by aberrations in the lineages of the cells that generate the vulva. Knowledge of the specific lineage defect in each strain has allowed us to assign 23 of these 28 genes to particular steps in the pathway of vulval development (Fig. 3). These assignments were confirmed by determining the epistatic interactions among mutations that affect different steps in the pathway and by analysing the temperature-sensitive periods of five of these genes.

Mutations in 19 of the genes that have been assigned to specific steps in this pathway cause certain cells to adopt fates characteristic of other cells. For example, mutations in three of the five genes necessary for the generation of the vulval precursor cells cause these cells to adopt fates characteristic of non-

precursor cells: the *lin-26* mutation causes the normally hypodermal precursor cells to adopt neural fates; the mutation *n300* causes the precursor cells to adopt an alternative hypodermal fate; and the *lin-4* mutation apparently causes the precursor cells in the L3 stage to adopt a fate like that of their own progenitors during the L1 stage. Mutations in the 15 genes (see Fig. 3) that affect the determination of the fates of the precursor cells cause certain precursor cells to adopt fates characteristic of other precursor cells. Mutations in *lin-11*, one of the three genes that affect the expression of the 2° lineage, cause certain 2°-specific cells to adopt fates of other 2°-specific cells.

We consider these specific transformations in cell fates to be homoeotic, by analogy with insect mutants in which one body part is replaced by another part normally found elsewhere in the animal^{17,18}. The existence of homoeotic vulval cell lineage mutants suggests that each step in the pathway of vulval development involves decisions that distinguish between alternative cell fates. Furthermore, there may be a regulatory hierarchy among the genes involved in these decisions. For example, *lin-12*, the activity of which is both necessary and sufficient for vulval precursor cells to become determined to express a 2° lineage, could act by regulating the activities of the genes involved in the expression of the 2° lineage, *lin-11*, *lin-17* and *lin-18*.

The nature of the transformations caused by these mutations suggests that we have identified genes involved in at least two possible mechanisms for specifying cell fates. First, mutations in *lin-11*, *lin-17*, *lin-18* and possibly *lin-26* disrupt or reverse the asymmetry of normally asymmetric cell divisions. If these divisions produce cells with autonomously determined fates, then these four genes may affect the generation of, segregation of, or response to determinants of cell fates that are normally asymmetrically distributed at certain cell divisions. Second, the 15 genes involved in the determination of the precursor cell fates act in the precursor cells in the reception of an intercellular signal and/or in the intracellular response to that signal. The discovery¹⁹ that one of these genes, *lin-12*, encodes a protein homologous both to the precursor of mammalian epidermal growth factor and to the low-density lipoprotein receptor supports the hypothesis that *lin-12* acts as a component of an intercellular signalling system. Other genes involved in the intracellular response to the anchor cell signal might encode membrane or cytoplasmic proteins that function in 'second messenger' systems similar to those that have been characterized biochemically from other organisms²⁰⁻²³ or nuclear proteins that function to control gene expression.

We thank Victor Ambros, Michael Bate, Martin Chalfie, Cynthia Kenyon, Eric Lander, Barbara Meyer, Aaron Mitchell, Gary Ruvkun, Dan Stinchcomb, Gary Struhl and members of our laboratory and of the laboratory of Barbara Meyer for valuable comments on this manuscript. This work was supported by USPHS research grants GM24663 and GM24943 and Research Career Development Award HD00369 to H.R.H. and by Predoctoral Training grant GM07287.

Received 15 September 1986; accepted 13 January 1987.

1. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110-156 (1977).
2. Deppe, U. et al. *Proc. natn. Acad. Sci. U.S.A.* **75**, 376-380 (1978).
3. Sulston, J. E., Albertson, D. G. & Thomson, J. N. *Dev. Biol.* **78**, 542-576 (1980).
4. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev. Biol.* **100**, 64-119 (1983).
5. Sulston, J. E. & White, J. G. *Dev. Biol.* **78**, 577-597 (1980).
6. Kimble, J. & Hirsh, D. *Dev. Biol.* **70**, 396-417 (1979).
7. Kimble, J. *Dev. Biol.* **87**, 286-300 (1981).
8. Sternberg, P. W. & Horvitz, H. R. *Cell* **44**, 761-772 (1986).
9. Horvitz, H. R. & Sulston, J. E. *Genetics* **96**, 435-454 (1980).
10. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **82**, 41-55 (1981).
11. Greenwald, I. S., Sternberg, P. W. & Horvitz, H. R. *Cell* **34**, 435-444 (1983).

12. Ferguson, E. L. & Horvitz, H. R. *Genetics* **110**, 17-72 (1985).
13. Fixsen, W. thesis, Mass. Inst. Technol. (1985).
14. Chalfie, M., Horvitz, H. R. & Sulston, J. E. *Cell* **24**, 59-69 (1981).
15. Ambros, V. & Horvitz, H. R. *Science* **226**, 409-416 (1984).
16. Sternberg, P. W. & Horvitz, H. R. *Dev. Biol.* **88**, 147-166 (1981).
17. Ouweneel, W. J. *Adv. Genet.* **18**, 179-248 (1976).
18. Garcia-Bellido, A. *Am. Zool.* **17**, 613-629 (1977).
19. Greenwald, I. *Cell* **43**, 583-590 (1985).
20. Schramm, M. & Selinger, Z. *Science* **225**, 1350-1356 (1984).
21. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
22. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
23. Nishizuka, Y. *Science* **225**, 1365-1370 (1984).
24. Suzuki, D. T. *Science* **170**, 695-706 (1970).
25. Hirsh, D. & Vanderslice, R. *Dev. Biol.* **49**, 220-235 (1976).

Is 3C324 the first gravitationally lensed giant galaxy?

O. Le Fèvre^{*†}, F. Hammer[†], L. Nottale[†] & G. Mathez[‡]

^{*} Canada-France-Hawaii Telescope Corporation, PO Box 1597, Kamuela, Hawaii 96743, USA

[†] DAEC, Observatoire de Meudon, 92195 Meudon Cedex, France

[‡] Observatoire du Pic du Midi et de Toulouse, 14 Av. Ed. Belin, 31400 Toulouse, France

The radiogalaxy 3C324 has recently raised considerable interest as one of the most distant (redshift $z = 1.206$) and luminous known galaxies. Here we report new observations of this object at Canada-France-Hawaii Telescope which confirm our previous predictions of the presence of an intervening gravitational lens. First, a multiple structure is revealed by our charge-coupled device (CCD) broad-band imagery. Second, following our previous identification in the 3C324 spectrum of an additional line system at $z = 0.84$, subsequent CCD observation through interferential filters were planned. In a wavelength range corresponding to an emission line at $z = 0.84$ only the central component is detected, while only outer components are seen in a range corresponding to a line redshifted by $z = 1.206$. These observations are interpreted as the first example of gravitational multiple imaging of a giant radiogalaxy by a foreground object. So galaxy 3C324 seems to be a gravitational mirage—though further spectroscopic observations will be needed to establish whether the two images have similar spectra.

On the basis of the confrontation of available data, with a theoretical analysis of the perturbations induced on the Hubble diagram by inhomogeneities in the actual Universe, two of us recently proposed that gravitational amplification of light affects far more extragalactic observations than currently believed¹⁻⁴. Extrapolation of this analysis to the Spinrad and Djorgovski⁵ very distant radiogalaxies led us to propose that their large observed overluminosity is at least partly due to strong gravitational lensing⁶. From case by case examination we have inferred the possibility of multiple imaging for some of them, including 3C324.

Based on our identification⁶ in the high quality spectrum published by Spinrad and Djorgovski⁷ of a line system (Ca II, H and K, and H β lines and MgH and MgB bands in absorption, [O II] in emission) due to a foreground galaxy at $z = 0.845 \pm 0.002$ superimposed on the narrow emission lines system at $z = 1.206$, new observations were planned.

Deep broad-band images were obtained by O.L. on 31 May and 1 June 1986 with the RCA2 CCD camera now available at Canada-France-Hawaii Telescope (CFHT)⁸. In spite of exposure times as long as 4,500 s in *R* (red wavelength) and 5,400 s in *I* (infrared wavelength), the full width at half maximum (FWHM) measured on stars is less than 0.7 arc s on both frames. In the *R* and *I* frames the multiple structure of 3C324 is evident⁹: two prominent components (a) and (b) separated by 1.1 arc s (12.9 h_{50}^{-1} kpc at $z = 1.206$, $q_0 = 0$) a fainter component (c) 0.8 arc s east of (a), and a probable faint component (d) 1.0 arc s north-west of (a) (Fig. 1a). Note also the two other galaxies (e) at 2.2 arc s from (b) and (f) at 3.0 arc s from (b). Calibrations on M92 field allowed us to derive the *R*–*I* colour index for each component: $R - I_{(a)} = 0.65 \pm 0.20$, $R - I_{(b)} = 1.05 \pm 0.20$, $R - I_{(c)} = 1.10 \pm 0.30$, $R - I_{(d)} = 1.07 \pm 0.40$. A comparison with the spectroscopic data of Spinrad and Djorgovski in the [O II] $\lambda 3727$ emission line at $z = 1.206$ (Fig. 1d) reveals that the (e) component, with a differential velocity around 1,000 km s⁻¹, is probably a galaxy at this redshift. Moreover, more than 100 galaxies are detected in only 1.15×1.86 arc min² confirming the hypothesis⁷ of a cluster around 3C324.

On 6 August 1986 a 3,600 s exposure was obtained in a 90-Å width interferential filter centred at 5,139 Å, very close to the $\lambda 2335$ [C II] emission line at $z = 1.206$ identified by Spinrad and Djorgovski. In this image (Fig. 1b) the measured seeing is 0.9 arc s: the components (b) and (c) only are detected while the brightest component in the broad-band image, (a), is obviously much fainter and probably non-emissive. The positions with respect to five stars in the field of both maxima in the *R* and 5,139 Å frames agree better than 0.13 arc s for (b) and 0.20 arc s for (c) (this larger error in the (c) position is due to the presence of the wings of (a) in the *R* frame).

On 7 and 8 August 1986 two 5,400 s exposures were obtained through an interferential filter with a 90 Å width centred at 6,840 Å, including [O II] $\lambda 3727$ at $z = 0.845$. The seeings measured on field stars are respectively 0.8 and 0.85 arc s FWHM. Figure 1c, the digital sum of both frames after proper recentering, shows component (a) now prominent, while neither (b) nor (c) are detected down to the level expected from broad-band photometry. Both positions of (a) in the *R* and 6,840 Å frames agree better than 0.12 arc s.

From these data we conclude that the (a) component is a bright foreground galaxy at $z = 0.845$, possibly a spiral from its spectrum, axis ratio ~ 0.4 and $R - I = 0.65$, while components (b), (c) and (e) are objects at $z = 1.2$. Note also that (a) is on the line joining (b) to (c). Consider some possible interpretations of such a configuration. At first sight, the proximity of components (b), (c) and (e) could agree with the suggestion by Spinrad and Djorgovski⁷ that some merging occurs in 3C324. But even if (b) and (c) were independent, the foreground galaxy (a) would act on (c) as a gravitational lens: from the condition of Subramanian and Cowlings¹⁰, it would inevitably give multiple imaging provided that its mass is greater than $1.5 \times 10^{11} M_{\odot}$ inside 8 h_{50}^{-1} kpc, a condition probably fulfilled because (a) is 10 times brighter than our own galaxy. Therefore we conclude that the gravitational lens hypothesis is the most probable.

It is outside the scope of this report to present detailed modelling of the 3C324 lensing configuration, which would handle the theory of multiple imaging of an extended source by an extended lens. However, we intend to show that available data may already be accounted for by a very simple model where not only components (b) and (c) but also possibly component (d) are images of a source galaxy at $z = 1.206$ lensed by galaxy (a) at $z = 0.845$.

Consider the Bourassa and Kantowski¹¹ theory and its simple application to spheroidal transparent gravitational lenses characterized by a cutoff radius a_m . The image configuration depends on the relative value of two parameters: $\mu = (4G/c^2)(M/a_m^2)$ and $\sin^2 \beta$. Here the usual distance term¹¹ D is equal to $0.074 c/H_0$ and $\cos \beta$, the observed deflector axis ratio, may be tentatively estimated to be ~ 0.45 from the $\lambda 6,840$ image. The facts that (b) and (c) components are dominant, lie on either sides and are aligned with the deflector, suggest that μ is large and that the lens behaves essentially as a point mass. Under this assumption^{12,13}, the image positions directly yield the critical radius $f = 10$ kpc and the source impact parameter $l = 4$ kpc, leading to a mass $M = 10^{12} M_{\odot}$, which is quite reasonable compared to the *R* absolute magnitude -21.5 . Consider now the complete Bourassa and Kantowski¹¹ theory: assuming $a_m = 10$ kpc for the observed deflector extension implies $\mu \sim 1 \sim \sin^2 \beta$. Depending on the line-of-sight position, one, three or five images may appear. There is very good agreement between the positions of the expected images 1 and 2 and of the components (b) and (c), accounting for the value previously derived for $l = 4$ kpc.

We may now compute the expected image amplifications in this model¹⁴. For component (b) and (c), we find respective amplifications $A_b = 1.80$ and $A_c = 1.0$. (Note that these values confirm the validity of our previous point mass approximation which would have given amplification factors of 1.82 and 0.82.)

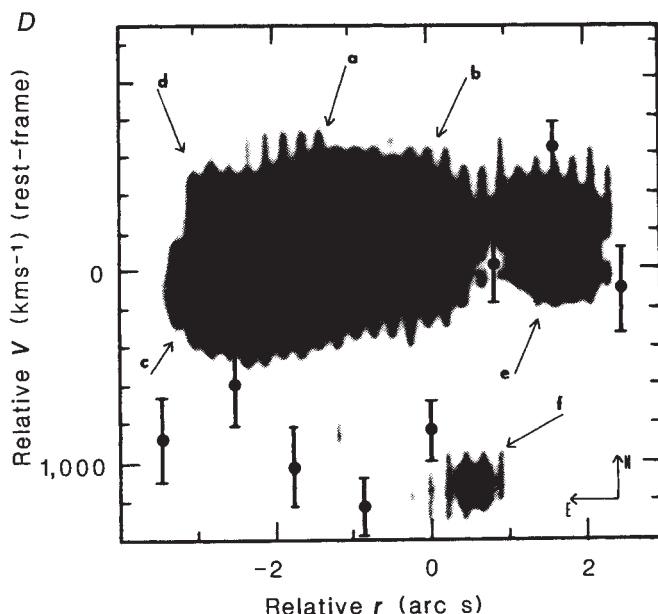
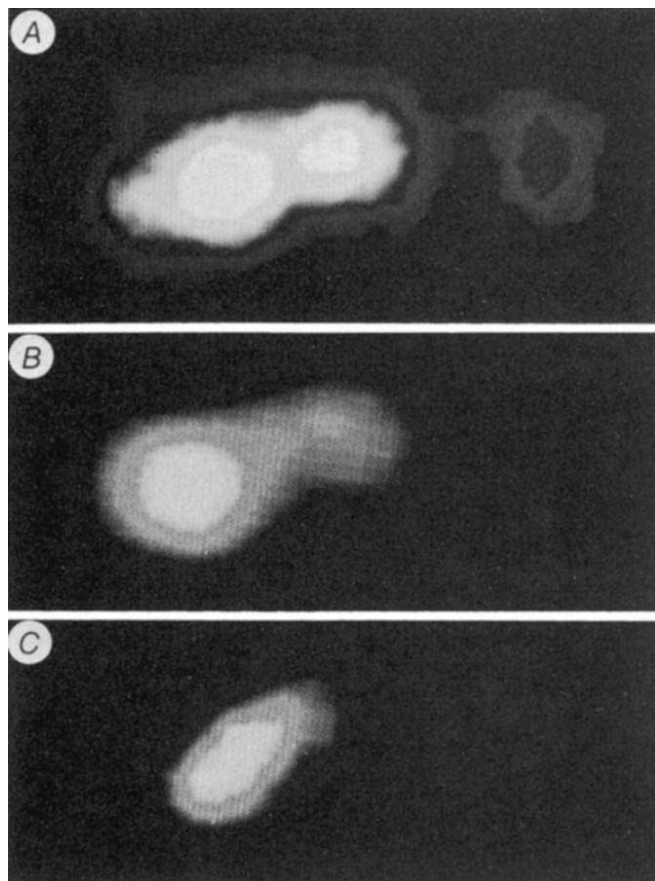


Fig. 1 Panel A, multiple structure of 3C324 in broad band R filter; B, imaging in a filter including the $[C\ II]$ emission line at $z = 1.206$. C, Imaging in a filter including the $[O\ II]$ emission line at $z = 0.845$. D, High contrast print of the R image of 3C324 best showing the faintest components (d), (e) and (f). Superimposed is the velocity field observed by Spinrad and Djorgovski⁷ from $[O\ II]$ emission at $z = 1.21$: there is no evidence for a non-zero differential velocity between (b) and (c) while (e) seems to differ by $\sim 1,000\text{ km s}^{-1}$ from (b) and (c).

The predicted amplifications agree well with R photometry which gives $R_b = 22.7$ and $R_c = 23.3$. As usual, the third expected image is so close to the deflector centre that overfocalization makes it unobservable.

Consider now component (d). It is quite remarkable that the observed axis ratio and position angle of deflector (a) have located it in the domain expected for images 4 and 5. Component (d) is, however, too faint for such an image position which predicts $d/c = 1.65$ instead of the observed value of 0.3. This discrepancy is easy to explain because, thanks to the source extension, a part of its outer region lies inside the diamond shaped figure of Bourassa and Kantowski¹¹ where five images are expected. Only 8% of the source light needs to fall within this limit to produce collapsed images 4 and 5 having the appearance of component (d). However, redshift information is clearly needed to confirm this configuration.

Other features of our data support our hypothesis, specifically the equality of the (b), (c) and (d) $R - I$ values (1.07), combined with their discrepancy from (a), $R - I$ (0.65), which agrees well with the gravitational lens interpretation of 3C324. Also, the model seems to exclude the presence of a rich galaxy cluster around the deflector at $z = 0.845$, which would drastically

increase the luminosity ratio b/c . As indicated by the redshift of component (e) galaxy E of Spinrad and Djorgovski⁵ and the high number of galaxies in the field, there is probably a cluster at $z = 1.2$.

In the $\lambda 5139$ imagery, component (c) appears brighter than (b), contrary to what is observed in the R frame. We find that image (c) in $\lambda 5139$ may be in fact a composite of (c), (d) and the continuum of (a) while in the $\lambda 6840$ filter the (b) and (c) expected continua are ~ 1 mag fainter than (a). Only observations in stronger lines like $[O\ II]\ \lambda 3727$ or $Mg\ II\ \lambda 2800$ at $z = 1.206$ would definitively clarify this point.

Finally, in agreement with our previous suggestions⁶, because of contamination by the foreground galaxy (a), and of total amplification by a factor of ~ 3 , it is clear that 3C324 is subjected to a luminosity excess of ~ 2 magnitudes, and a colour excess of ~ 0.2 magnitudes with respect to its real magnitude and $R - I$ colour index.

We thank C. Christian for help during telescope observations and B. Fort, Y. Mellier and G. Soucail for valuable suggestions. We also thank Dr H. Spinrad for useful discussions. The work presented here is based on observations made using the Canada-France-Hawaii Telescope.

Received 30 October; accepted 30 December 1986.

1. Nottale, L. & Hammer, F. *Astr. Astrophys.* **141**, 144-150 (1984).
2. Hammer, F. *Astr. Astrophys.* **152**, 262-270 (1985).
3. Hammer, F. & Nottale, L. *Astr. Astrophys.* **155**, 420-422 (1986).
4. Hammer, F. & Nottale, L. *Astr. Astrophys.* **167**, 1-10 (1986).
5. Spinrad, H. & Djorgovski, S. *Astrophys. J.* **285**, L49-L52 (1984).
6. Hammer, F., Nottale, L. & Le Fèvre, O. *Astr. Astrophys.* **169**, L1-L3 (1986).
7. Spinrad, H. & Djorgovski, S. *Astrophys. J.* **280**, L9-L12 (1984).

8. Christian, C., Waddell, P., McGonegal, R. *SPEI Proc.* no. 627, 500-508 (1986).
9. Le Fèvre, O. *IAU Telegram* No. 4233 (1986).
10. Subramanian, K. & Cowlings, S. A. *Mon. Not. R. astr. Soc.* **219**, 333-346 (1986).
11. Bourassa, R. R. & Kantowski, R. *Astrophys. J.* **195**, 13-21 (1975).
12. Refsdal, S. *Mon. Not. R. astr. Soc.* **128**, 295 (1964).
13. Nottale, L. *Astr. Astrophys.* **157**, 383-386 (1986).
14. Young, P. *Astrophys. J.* **244**, 756-767 (1981).
15. Cawthorne, T. V., Scheuer, P. A. G., Morijon, I. & Muxlow, T. W. B. *Mon. Not. R. astr. Soc.* **219**, 883-893 (1986).

The optical polarization of the Sun measured at a sensitivity of parts in ten million

J. C. Kemp*, G. D. Henson*, C. T. Steiner*† & E. R. Powell†

* Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

† Central Oregon Community College, Bend, Oregon 97701, USA

The integrated light of the Sun, an essentially spherical star with only slight asymmetries (small oblateness, weak overall magnetic field), would normally be found to be unpolarized, as observed in broad spectral bands with common instrumental sensitivities, 10^{-4} – 10^{-5} fractional polarization^{1–4}. Defining the Sun's intrinsic linear (LP) and circular (CP) polarizations down to much lower levels (10^{-7} or even 10^{-8}) would have consequences not only in solar physics but in other areas, setting, for example, a new standard for stellar polarimetry. We report measurements of the general polarization of the Sun, both CP and LP, over the whole disk and over large sectors, at an absolute sensitivity level of $\leq 3 \times 10^{-7}$, carried out during August–September 1986. Upper limits for the intrinsic whole-Sun LP from the best data (minimum Earth-atmospheric contamination) were 0.2×10^{-6} in the V (yellow) band and 0.8×10^{-6} in B (blue). Definite CP was discovered. (1) The north and south polar zones showed values of V (normalized or fractional CP) of -1 to -6×10^{-6} for the north, and 0 to $+2 \times 10^{-6}$ for the south zones. The spectral CP rises steeply toward the blue. (2) The whole disk had a net CP of -0.1 to -1.0×10^{-6} (from red to blue), negative as with the magnetically dominant north pole. The spectral dependence of the global broadband CP resembles that of sunspots^{5,6} and of local non-spot regions with magnetic flux tubes⁷.

An unusual arrangement was used for this high-sensitivity polarimetry of the whole Sun and large portions of its disk, shown in Fig. 1. The key element, a photoelastic modulator^{3,8} (PEM) (manufactured by Hinds International, Portland, Oregon)—or sometimes a pair of PEMs—was mounted on the upper rim of the 81-cm telescope at Pine Mountain Observatory, Oregon. The telescope itself was used as a solar spar, with its main aperture blocked off by a shroud. The PEM faced the Sun directly, with no intervening optics which might cause spurious polarization, and below it were mounted an upper aperture plate; an optional lens; a stepping-motor rotated Polaroid analyser; a lower, image-plane aperture for use with the lens; a filter; and the detector, a blue-enhanced, 1 cm-diameter, flat-faced silicon photodiode, with associated preamplifier. The PEMs were fused-silica isometric types³. One or both of two units, of frequencies 32.4 and 35.1 kHz, were used. Detection and processing of the modulated polarized-light signals from the photodetector were done by standard electronics including a lock-in detector and computer². Several spurious signals were cancelled by the rotating analyser (second stage chopper^{2,8}) and by modulator rotation⁸. All data presented here are those averaged over four cardinal orientations of the analyser; and over four or eight orientations of the PEM or PEMs, using a rotator plate (Fig. 1).

For absolute polarimetry at levels $< 10^{-6}$, certain instrumental effects not previously considered must be dealt with. In all polarimeters which employ birefringence modulators, including PEMs, Pockels cells and rotating wave plates, a fundamental aberration is caused by dichroic surface transmittance, which is inherent in Fresnel's laws of refraction at a dielectric interface. This imposes an oscillating linear polarization on the incident light, at the surfaces of the modulating device. In circular polarimetry, specifically, the CP is usually converted into a

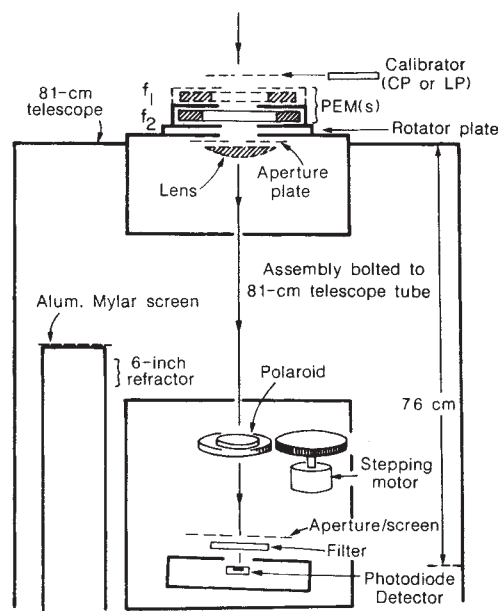


Fig. 1 Solar polarimeter assembly attached to the 81-cm telescope at Pine Mountain Observatory, Oregon. Telescope was used only as a solar spar, with entrance covered. Guiding was done with the 6-inch guide telescope, and by directly viewing the image on the detector. Apparatus was accessed from a scaffold erected by the telescope. All measurements involved successive measures with the photoelastic modulator (PEM) assembly rotated to four or eight cardinal positions around the system axis, and with the polaroid analyser rotated to four positions, to cancel out various instrumental signals. Lens was used for sector measurements (Fig. 2); no lens was used in the whole-disk observations.

modulated LP by the device, and the LP is then converted into an intensity modulation by a stationary analyser-polarizer. The surface-dichroism effect produces a spurious, similar modulation. The false CP so induced can be as large as 10^{-4} . Because the linear polarizations produced by the real CP and the surface dichroism are 45° apart around the system axis, in principle the false signal can be eliminated by precise alignment of the analyser at 45° from the modulator eigen-axes. Owing to device imperfections this false signal cannot be cancelled with confidence by more than a factor of about 100, merely by device alignment. In our apparatus we have, instead, eliminated the false surface-dichroism signal by arranging for CP detection not at the basic modulator frequency, at which the false signal occurs, but at a different frequency. In one method, we employed two PEMs in tandem, operating at frequencies f_1 and f_2 , with retardation amplitudes $\lambda/2$ and $\lambda/4$, respectively. This combination, together with the stationary analyser, produces a heterodyne CP signal in the detector of frequency $2f_1 - f_2$; in our case this was 29.7 kHz. In a second method we used one PEM in a third-harmonic mode, with the retardation amplitude set at essentially $3\lambda/4$ and a third-harmonic CP signal (frequency $3f_1$) was generated. Some of the CP data reported here, namely the data of 7–8 August in Fig. 2 and 27 July in Fig. 3, were obtained with the heterodyne ($2f_1 - f_2$) method; all the other data in Figs 2–3 were obtained with the $3f$ method. No systematic difference is seen between the results with these different arrangements.

Linear polarimetry was performed by the second-harmonic method⁹, which requires rotating a single PEM to two positions, 45° apart, to obtain the two linear Stokes parameters Q and U . In fact we used eight cardinal orientations, with suitable averaging. In the LP case, second-harmonic detection immediately eliminates false signals due to the surface-dichroism effect discussed above. In our apparatus the only instrumental LP was judged to be caused by slight misalignment of the system from

† Present address: Department of Physics, Boston University, Boston, Massachusetts 02215, USA.

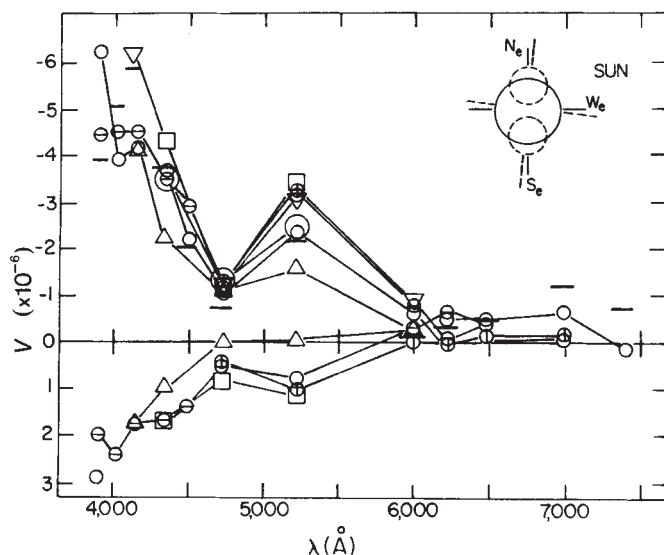


Fig. 2 Filter-band measures of the Sun's circular polarization, for the north pole, *a*, and the south pole, *b*, using a 15 arc min-diameter aperture at the focal plane over the photodetector. Approximate positions on the Sun's disk are sketched at upper right; positions were over the ecliptic poles, N_e and S_e , which were rotated slightly from terrestrial north and south (which are indicated by broken lines) during the observing period. Horizontal line segments (without symbols), approximate CP values of a typical sunspot, extracted from data in ref. 5, for interference filter bands equivalent to those used in this work; these values are scaled arbitrarily on the vertical scale for visual comparison with the north pole data. The ordinate scale here is in units of 10^{-6} fractional polarization. Filter widths used were from 60–200 Å. For brevity, in Figs 2 and 3 and in the text, V denotes the normalized quantity V/I , in terms of Stokes parameters. $\circ$, 7 August; Φ , 10 August; $\bigcirc$, 15 August; ∇ , 2 September; $\ominus$, 8 August; $\square$, 11 August; Δ , 22–23 August (1986).

collinearity with the light axis. The spurious LP is of the order of $0.1 \theta^2$, where θ is the misalignment angle of the PEM surface³. In our case we estimated that $\theta \leq 0.1^\circ = 0.0017$ rad, implying that $p_{\text{inst}} \leq 3 \times 10^{-7}$. As the Sun is an extended object, this estimate applies most specifically to whole-disk observations with the Sun centred in the system aperture.

Many tests and modifications of the apparatus were made during early phases of this project, for example, the elimination of inter-reflections, which we accomplished by slightly tilting the analyser and filter-detector assemblies (Fig. 1). At a suggestion by J. Stenflo (personal communication) we experimented with an auxiliary 'scrambler' PEM¹⁰, placed above the main PEM, which in theory can eliminate the real signal and reveal spurious ones. We found that the scrambler itself introduced new false signals. Our general conclusion is that no optical components of any kind should intervene between the observed source (the Sun) and the principal polarimeter element. No light source known to be unpolarized at the 10^{-7} level was available to test this system. Nevertheless, based on intimate knowledge of the apparatus and on consistencies in the final data, we believe that the instrumental levels in most of the CP and LP data here reported (Figs 2–4) do not exceed $\sim 3 \times 10^{-7}$. Even greater sensitivity is possible in future work.

In Figs 2 and 3 we show north and south and global measures of the Sun's broadband CP, given in terms of the normalized polarization (circularly polarized flux divided by the total flux). The north and south polar measurements were made with an aperture diameter of 15 arc min, with approximate positions as shown in Fig. 2. In the whole-disk measures, Fig. 3, the effective aperture was 2° . Interference filters of various widths, 60–200 Å, were used. The north-south measurements were emphasized partly because the north-south dipolar magnetic field is strong

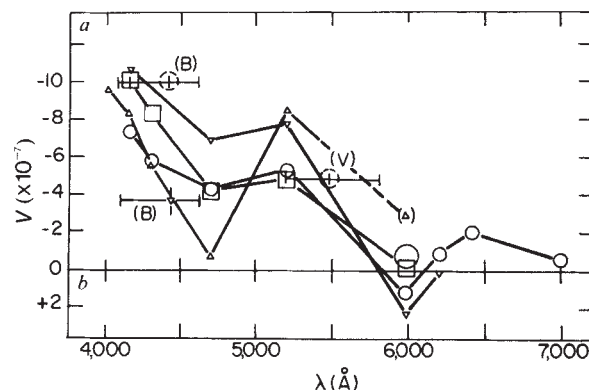


Fig. 3 The circular polarization of the Sun's whole disk. No lens was used, and an angular aperture of diameter 2° was defined only by the upper aperture plate in Fig. 1. Although we do not show individual measurement errors, the formal errors merely from photon statistics are mostly small compared to the scatter between points from different dates, caused by instrumental polarization changes and possibly changes in the Sun's polarization. Note that the vertical (fractional CP) scale here is in units of 10^{-7} , rather than 10^{-6} as in Fig. 2. The measurements were taken on: $\circ$, 27 July; Δ , 16 August; $\bigcirc$, 19 August; $\square$, 21 August; ∇ , 24 August; $\ominus$, 2 September (1986).

at this time in the solar cycle, close to sunspot minimum. A few east-west measurements were taken, using the aperture as in Fig. 2 rotated to the east and west limbs, all in blue filter bands (see Table 1). No sunspots were evident in the north-south measurements. The few spots seen during the series were near the equator, as is typical near sunspot minimum. One or two of the east-west measurements may have been influenced by spots. The CP quantity V in Figs 2 and 3, and in Table 1, is the normalized circular Stokes parameter (divided by the intensity I). For the signs of V we use the helicity convention: positive V means that, in a stationary plane in front of the observer, the E vector is seen to rotate counter-clockwise.

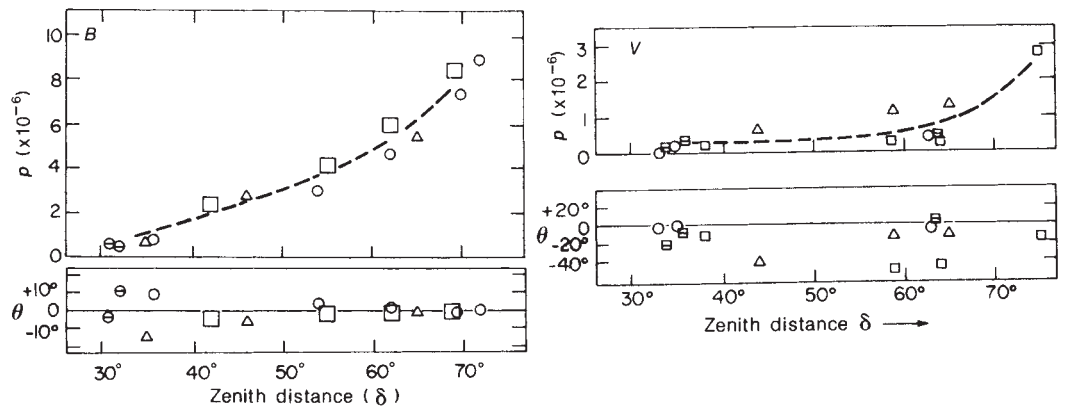
In Fig. 1 note: (1) a marked wavelength dependence, including a general rise into the blue ultraviolet and a dip around 4,700 Å. Such structure, especially the 4,700-Å drop, cannot be of instrumental origin and is proof that we are detecting real polarization; (2) a dominance of the Sun's CP by the north sector. Regarding the spectral dependence we include in Fig. 2 horizontal line segments showing equivalent sunspot CP values, in filter bands comparable to those used in the present sector measures, extracted from Fig. 2 of Kemp and Henson⁵. An arbitrary sign and an arbitrary scale are used for the sunspot CP values, which are ~ 100 times larger than the north-south sector values. The similarity in spectrum is striking. In the whole-disk data of Fig. 3, a basically similar wavelength dependence is seen, although the values are 5–10 times smaller than in the north aperture of Fig. 2, and the data show considerable scatter. Clearly the negative sign of the whole-disk values reflects the dominance of the north

Table 1 East-west circular polarization measures of the sun

Date 1986	Filter centre (Å)	V (east) ($\times 10^{-6}$)	V (west) ($\times 10^{-6}$)	$(V_E - V_W)/2$ ($\times 10^{-6}$)
2 August	4,300 (B)	-0.17	+0.48	-0.32
3 August	4,300 (B)	-2.76	+1.84	-2.30
5 August	4,300 (B)	-2.46	+1.41	-1.94
11 August	4,306	-2.14	+0.05	-1.09
23 August	4,150	-0.26	-3.01	+1.40
24 August	4,150	+0.89	-2.11	+1.50
2 September	4,150	+1.46	+0.34	+0.56

15 arc min-diameter aperture, positions analogous to north-south positions in Fig. 2.

Fig. 4 Measured values of the Sun's whole-disk linear polarization, obtained as in the CP measurements of Fig. 3 without imaging (no lens), using a 2° aperture. Standard astronomical *B* and *V* filters (mean wavelengths 4,400 Å and 5,530 Å respectively) were used. The vertical scales for *p* (polarization amplitude) are in units of 10⁻⁶ fractional LP. We show *p*(δ) and θ' (δ), where δ is the zenith distance (Sun-zenith angle) and $\theta' = \theta - \theta_{\text{zen}}$ is the difference between the position angle of the Sun's observed polarization in the equatorial system, and the Sun-zenith direction on the sky. At large zenith angles the observed LP, at least in the blue band, is almost entirely controlled by atmospheric scattering. In the *V* band, owing to the very small *p* values the position angles were very poorly defined. For the *B* filter diagram the measurements were taken on: $\ominus$, 12 August (a.m.); $\circ$, 12 August (p.m.); Δ , 25 August (a.m.); $\square$, 26 August (a.m.); (1986). For the *V* filter diagram the measurements were taken on: $\circ$, 10 August (a.m.); $\boxplus$, 12 August (a.m.); $\square$, 12 August (p.m.); Δ , 26 August (a.m.) (1986).



polar CP. Note that these data span roughly a rotation period of the Sun, so that azimuthal structure can be largely averaged out. In these CP measurements there was no indication of Earth-atmospheric influences, such as zenith-distance variation.

Our linear polarization results, in contrast to the CP measures, were controlled mainly by terrestrial atmospheric effects. Only whole-disk measures were taken, in 1°- and 2°-diameter apertures, in the *B* and *V* bands. The observed LP correlated in magnitude and position angle with the Sun's position on the sky, at least in the *B* band wherein the LP was strongest and better defined. In Fig. 4 the LP is shown in terms of the magnitude *p* (in units of 10⁻⁶ fractional LP); and a position angle $\theta' = \theta - \theta_{\text{zen}}$, where θ is the actual position angle (E vector direction) in the standard equatorial system and θ_{zen} is the position angle of the Sun-zenith direction. Thus $\theta' = 0$ means that the E vector is along the Sun-zenith great circle. Such LP is termed 'negative' in the terminology of terrestrial sky polarization¹¹. This negative LP, increasing with zenith distance, is attributed to double scattering by aerosols and molecules in the Earth's atmosphere¹¹. Using a 1° aperture we also found that the adjacent sky around the Sun, in the so-called aureole¹², had comparable LP flux as in the Sun measures, with the same position angle. We conclude from Fig. 4 that the Sun's intrinsic whole-disk LP does not exceed the levels at the smallest zenith angles, that is, about 0.8×10^{-6} in *B* and 0.2×10^{-6} in *V*. Better LP measures should be taken at the zenith (at a low-latitude site) or in space.

The polar circular polarization, with opposing signs north and south, is probably a manifestation of the Sun's dipolar magnetic field, which is approximately at maximum strength now (sunspot minimum). A referee has noted that the northern and southern hemispheres had different magnetic structures during this period, the northern hemisphere having a smooth appearance whereas the southern hemisphere was magnetically complex. In any case, however, the northern zone during 1986 was magnetically negative, as shown for example by the Stanford magnetograph measures (ref. 13 and 1986 data supplied by P. Scherrer and J. Hoeksema), whereas the south pole was positive. Thus the sign relationship here is that our negative broadband *V* in the northern zone corresponds to a predominant negative (inward) field, and vice versa in the south. This sign correlation will be one test of theories of the global CP.

Regarding the origin of the global CP, as noted above the coarse spectrum (Figs 2 and 3) is rather similar to the broadband CP spectrum of sunspots⁵ and also to the CP spectrum of mildly magnetic regions away from spots, containing flux tubes (Fig. 5 of ref. 7). Thus a common mechanism for the CP is indicated. The coarse *V*(λ) spectra reflect the photospheric line blanketing,

which generally rises toward the blue, apart from local (essentially statistical) irregularities. The source of the broadband CP obviously lies in the spectral lines. Kemp and Henson⁵ showed that in sunspots the continuum contribution, from grey magnetopacity, is very weak in visible bands and is significant only in the infrared. In the case of broadband CP observed in flux-tube regions away from spots, it has been shown⁷ that this is an aspect of asymmetries between the blue and red peaks of the *V*(λ) curves of individual lines^{14,15}, which have opposite polarities but systematically different total strengths. The most quoted candidate source for this asymmetry, and for the resultant broadband CP, is mass motion in the flux tubes. Correlated vertical field and velocity gradients, causing differing Zeeman and Doppler shifts at different depths, combine with line saturation to produce the asymmetry¹⁶. At present there are several problems with quantitative models of this as discussed recently by Stenflo *et al.*¹⁷.

The Sun's surface is known to be threaded by a network of magnetic flux tubes, of sub-arc s diameters, with generally alternating polarities (see, for example ref. 18). Large-sector and global broadband CP must arise, in a manner of speaking, from a slight excess of tubes of one polarity, which account for an average vector field component in the observer's direction. Regarding the magnitudes of our observed global CP, the vector line-of-sight field in our polar-sector measures (Fig. 2) can be gauged from magnetograph data¹³ as perhaps being some 2 gauss in either polar region. From sunspot *V* values near 4,100 Å of $\sim 6 \times 10^{-4}$ with 2 kilogauss fields⁵, we might anticipate $V \sim 1 \times 10^{-6}$ in our north or south apertures, which is a few times smaller than we observe around this wavelength (Fig. 2). On the other hand a rough scaling based on typical Zeeman line asymmetries observed in non-spot regions (J. W. Harvey, personal communication), assuming reasonable area filling factors for the flux tubes, seems consistent with a *V* of a few parts in 10^6 as we observe.

The signs of our measured *V* values, with given polarity of the vector line-of-sight fields, should be compared with the signs expected from Zeeman line-wing asymmetries in magnetic non-spot regions^{7,14}. In all areas well away from the limb, the latter predict a negative broadband *V* in a region with positive (outward) field, and conversely. (We adhere to the helicity sign convention defined above.) In our north-south measures (Fig. 2), we see just the opposite. However, it has been recently found that near the solar limb, for $\mu = \cos \theta \leq 0.4$ where θ is the angle between the surface normal and the observer's direction, the line-wing area asymmetries change sign¹⁷. Evidently, in our north-south apertures the reversed 'limb' sign predominates. The polarized-light radiative transfer in flux tubes on

inclined surfaces is clearly complex, and more is involved than merely the line-of-sight field and its polarity¹⁷. In spite of this complexity, there can be no doubt that the broadband circular polarization mirrors, semi-quantitatively, the magnetic field structure, as has been suggested⁷.

The whole-disk CP of Fig. 3 might reflect only a transient condition of the Sun in which the northern hemisphere was dominated by a relatively ordered field. Or, it might be caused by the tilt of the Sun's spin axis combined with the north-south dipolar field. On about 8 September 1986 the Sun's north pole was inclined at its maximum angle (7.2°) towards the Earth¹⁹. Throughout August it was inclined at least 5.7° towards the Earth, which could explain the dominance of the north polar CP. Whole-disk magnetographs during August 1986 (J. T. Hoeksema and P. H. Scherrer, personal communication) confirmed an average negative magnetic polarity for the period. It is not clear how the rather small tilt angle of the Sun's axis could produce such a sizeable whole-disk CP and such a large difference in the north and south polar measures. Observations over years will be needed to investigate this.

A net circular polarization in the Sun's light, averaged over the whole visible spectrum, has implications in areas other than solar physics. In the solar-terrestrial connection, we wonder whether a quite small circular polarization might have influenced the chirality (average handedness) of the Earth's biomolecules. In the light received by the Earth, the Sun's overall CP would tend to average to zero over long periods, owing to both the Earth's orbital motion around the Sun's tilted spin axis and the reversal, every 11 years, of the Sun's dipolar field. On timescales of 10⁹ years, there would seem to be no net effect. However, if biogenesis began not slowly but in a burst of photochemistry, over a period of months perhaps, then a net CP in sunlight might have promoted chirality. It is not clear whether a CP as small as 10⁻⁶ could have been effective in this connection. The structure, strength and periodicity of the Sun's global magnetic field might have been quite different 3 Gyr ago.

Magnetic fields in solar and late-type stars, associated with star spots and related phenomena, may now be profitably looked for by the broadband CP method, presuming that the same process underlying the Sun's global magnetic CP operates in similar stars. Searches for magnetic fields in late-type stars have largely used indirect, non-polarimetric methods such as the sensing of Zeeman broadening of spectral lines^{20,21}. Scaling from our solar results, we note that a 10-gauss line-of-sight vector field in such stars would produce a broadband CP of possibly 10⁻⁵, which we have shown to be readily detectable. In the case of one of the brightest of these stars, the star Lambda Andromedae, for which there is strong evidence for Zeeman line broadening²¹, we will soon submit (Kemp *et al.*, in preparation) a report on the direct detection of broadband circular polarization.

This unusual project and the interpretation benefited greatly from the encouragement and advice of colleagues including Jack Harvey, A. Skumanich, Jan Stenflo, Phil Scherrer and S. Solanki. Ted Kinsman aided in some of the measurements. We received support from NSF grant ATM 841176 and the National Geographic Society grant 3003-84.

Note added in proof: We call attention to an important new paper on the theory of broadband polarization in the Sun and sunspots by A. Skumanich and B. W. Lites, *Astrophys. J.* (submitted).

Received 29 September; accepted 15 December 1986.

1. Dollfus, A. in *Planets, Stars and Nebulae Studied with Photopolarimetry* (ed. Gehrels, T.) 695-729 (Univ. Arizona Press, Tucson, 1974).
2. Kemp, J. C. & Barbour, M. S. *Publ. astr. Soc. Pacific* **93**, 521-525 (1981).
3. Kemp, J. C. *Proc. Soc. photo-opt. Instrum. Eng.* **307**, 83-88 (1981).
4. Tinbergen, J. *Astr. Astrophys.* **105**, 53-64 (1982).
5. Kemp, J. C. & Henson, G. D. *Astrophys. J.* **266**, L69-L72 (1983).
6. Henson, G. D. & Kemp, J. C. *Sol. Phys.* **93**, 289-299 (1984).
7. Stenflo, J. O. *Appl. Opt.* **23**, 1267-1278 (1984).
8. Kemp, J. C., Wolstencroft, R. D. & Swedlund, J. B. *Astrophys. J.* **177**, 177-189 (1972).
9. Stokes, R. A., Ekstrom, P. A. & Swedlund, J. B. *Opt. Eng.* **15**, 7-11 (1976).
10. Cheng, J. C., Nafie, L. A. & Stephens, P. J. *J. opt. Soc. Am.* **65**, 1031-1035 (1975).

11. Coulson, K. L. in *Planets, Stars and Nebulae Studied with Photopolarimetry* (ed. Gehrels, T.) 444-471 (Univ. Arizona Press, Tucson, 1974).
12. Pierce, A. K. *Astrophys. J.* **120**, 221-232 (1954).
13. Hoeksema, J. T. & Scherrer, P. H. *The Solar Magnetic Field—1976 through 1985* (Report UAG-94, Natn. Geophysical Data Center, Boulder, USA, 1976).
14. Stenflo, J. O., Harvey, J. W., Brault, J. W. & Solanki, S. *Astr. Astrophys.* **131**, 333-346 (1984).
15. Stenflo, J. O. & Harvey, J. W. *Sol. Phys.* **95**, 99-118 (1985).
16. Auer, L. H. & Heasley, J. M. *Astr. Astrophys.* **64**, 67-71 (1978).
17. Stenflo, J. O., Solanki, S. K. & Harvey, J. W. *Astr. Astrophys.* (in the press).
18. Stenflo, J. O. in *Basic Mechanism of Solar Activity* (eds Bumba, V. & Kleczek, W.) 69-99 (Reidel, Dordrecht, 1976).
19. Roberts, C. K. & Boksenberg, A. (eds) *The Astronomical Almanac 1986 C15* (US Govt Printing Office and HMSO, 1986).
20. Mullan, D. J. *Astrophys. J.* **279**, 746-747 (1984).
21. Giampapa, M. S., Golub, L. & Worden, S. P. *Astrophys. J.* **268**, L121-L125 (1983).

Evidence for two intervals of enhanced ¹⁰Be deposition in Antarctic ice during the last glacial period

G. M. Raisbeck*, F. Yiou*, D. Bourles*,
C. Lorius†, J. Jouzel‡ & N. I. Barkov§

* Laboratoire Rene Bernas, 91406 Orsay, France

† Laboratoire de Glaciologie et Geophysique de l'Environnement, 38402 Saint Martin d'Heres, France

‡ Laboratoire de Geochimie Isotopique—LODYC (UA CNRS 1206), CEN Saclay, 91190 Gif-Sur-Yvette, France

§ Arctic and Antarctic Research Institute, Fontanka 34, Leningrad 191104, USSR

We have previously reported concentration profiles of cosmic ray produced (cosmogenic) ¹⁰Be in deep ice cores from Dome C and Vostok, Antarctica^{1,2}. In both these cores we found a concentration of ¹⁰Be approximately two times larger in ice from the late glacial period than in the Holocene ice. We interpreted this as probably resulting from a lower precipitation rate on the Antarctic plateau during glacial periods, compared to interglacial periods. In the Vostok profile there was one sample, corresponding to ~60,000 yr BP, which gave an unusually large ¹⁰Be concentration, not correlated with any obvious climatic event. We suggested that this sample might be reflecting increased ¹⁰Be production, as for example during a period of reduced solar modulation². We have now measured a much more detailed concentration profile for ¹⁰Be in the Vostok core. The results confirm a ¹⁰Be 'peak', lasting ~1,000-2,000 years, at ~60,000 yr BP, and show another similar peak at ~35,000 yr BP. We have also observed the latter peak in the Dome C core. We discuss possible sources for these peaks, and their potential as stratigraphic markers.

Our previous Vostok profile was based on samples from 22 different depth levels². We now have measurements for 92 levels, including almost all those used for the oxygen isotope profile of ref. 3. In addition we have subdivided the samples from a number of the levels (including the two 'peaks') into two or more fractions, to obtain some information on short-term fluctuations in the ¹⁰Be concentration. The accelerator mass spectrometry measurements of ¹⁰Be were carried out as described in ref. 2, with the exception that in this work only 0.25 mg of ⁹Be carrier was used. For virtually all the new samples, 1,000 or more ¹⁰Be nuclei were recorded, making the one standard deviation statistical contribution to the measurement uncertainty <3%. Better stability in our operating conditions has also allowed us to reduce from 7 to 5% the instrumental contribution to our estimated uncertainties for most of the new measurements.

The multiple measurements for those samples subdivided, together with the length and estimated deposition time of the individual subsamples, are summarized in Table 1. It can be seen that in most cases the concentrations in the subsamples agree to within 10% of that for the mean concentration at that level. There are, however, exceptions in which the variability is considerably larger, perhaps due to short-term precipitation

Table 1 ^{10}Be concentration in levels subsampled

Sample designation	Nominal depth* (m)	Sample length (cm)	Estimated deposition interval† (yr)	^{10}Be conc. ($\times 10^3$ atoms g^{-1})	Mean ^{10}Be for level ($\times 10^3$ atoms g^{-1})
1c	25	18	7	97 $\pm$ 5	89 $\pm$ 5
1d	25	15	6	79 $\pm$ 4	
10f (1)	250	11	5	87 $\pm$ 7	86 $\pm$ 6
(2)	250	11	5	86 $\pm$ 5	
12c (1)	300	12	6	100 $\pm$ 6	113 $\pm$ 6
(2)	300	12	6	125 $\pm$ 7	
18b (1)	450	9	8	211 $\pm$ 11	218 $\pm$ 12
(2)	450	9	8	227 $\pm$ 12	
19a (1)	475	13	11	162 $\pm$ 8	180 $\pm$ 9
(2)	475	13	11	167 $\pm$ 9	
(3)	475	13	11	204 $\pm$ 10	
23d (1)	575	7	6	162 $\pm$ 8	193 $\pm$ 10
(2)	575	7	6	176 $\pm$ 9	
(3)	575	7	6	243 $\pm$ 12	
24a (1)	600	12	10	346 $\pm$ 18	316 $\pm$ 16
(2)	600	12	10	283 $\pm$ 15	
26d (1)	650	8	7	221 $\pm$ 12	208 $\pm$ 11
(2)	650	8	7	196 $\pm$ 11	
29b	725	14	11	156 $\pm$ 8	152 $\pm$ 8
29d	725	11	8	146 $\pm$ 8	
37a	925	25	21	266 $\pm$ 19	269 $\pm$ 20
37b (1)	925	7	6	212 $\pm$ 17	
(2)	925	6	5	231 $\pm$ 21	
(3)	925	7	6	302 $\pm$ 21	
(4)	925	6	5	342 $\pm$ 27	
38e (1)	950	7	6	136 $\pm$ 10	139 $\pm$ 10
(2)	950	7	6	141 $\pm$ 11	
(3)	950	9	8	141 $\pm$ 11	
40b (1)	1,000	8	7	180 $\pm$ 10	171 $\pm$ 9
(2)	1,000	8	7	162 $\pm$ 9	
42a (1)	1,050	11	10	214 $\pm$ 11	196 $\pm$ 10
(2)	1,050	11	10	175 $\pm$ 9	
43a (1)	1,075	10	9	113 $\pm$ 6	118 $\pm$ 7
(2)	1,075	10	9	125 $\pm$ 9	
44b (1)	1,100	9	7	114 $\pm$ 6	134 $\pm$ 7
(2)	1,100	9	7	154 $\pm$ 8	
48b (1)	1,200	10	8	114 $\pm$ 6	114 $\pm$ 6
(2)	1,200	10	8	114 $\pm$ 6	
54a (1)	1,350	15	12	167 $\pm$ 9	162 $\pm$ 9
(2)	1,350	15	12	157 $\pm$ 9	
62d (1)	1,550	21	20	132 $\pm$ 7	140 $\pm$ 8
(2)	1,550	19	18	152 $\pm$ 8	
63c (1)	1,575	10	10	142 $\pm$ 8	119 $\pm$ 7
(2)	1,575	10	10	93 $\pm$ 5	
65a (1)	1,625	10	9	116 $\pm$ 6	125 $\pm$ 7
(2)	1,625	10	9	134 $\pm$ 7	
67f	1,675	10	7	68 $\pm$ 4	72 $\pm$ 4
67j	1,675	10	7	77 $\pm$ 4	
69a	1,725	15	11	68 $\pm$ 4	72 $\pm$ 4
69d	1,725	10	7	77 $\pm$ 4	
70b (1)	1,750	16	12	64 $\pm$ 4	70 $\pm$ 4
(2)	1,750	16	12	77 $\pm$ 4	
70c	1,750	20	15	72 $\pm$ 6	
71f (1)	1,775	11	8	74 $\pm$ 4	76 $\pm$ 4
(2)	1,775	11	8	78 $\pm$ 4	
81b (1)	2,025	12	21	98 $\pm$ 5	104 $\pm$ 5
(2)	2,025	14	25	110 $\pm$ 6	
92c (1)	2,050	20	35	123 $\pm$ 6	127 $\pm$ 7
(2)	2,050	20	35	131 $\pm$ 8	

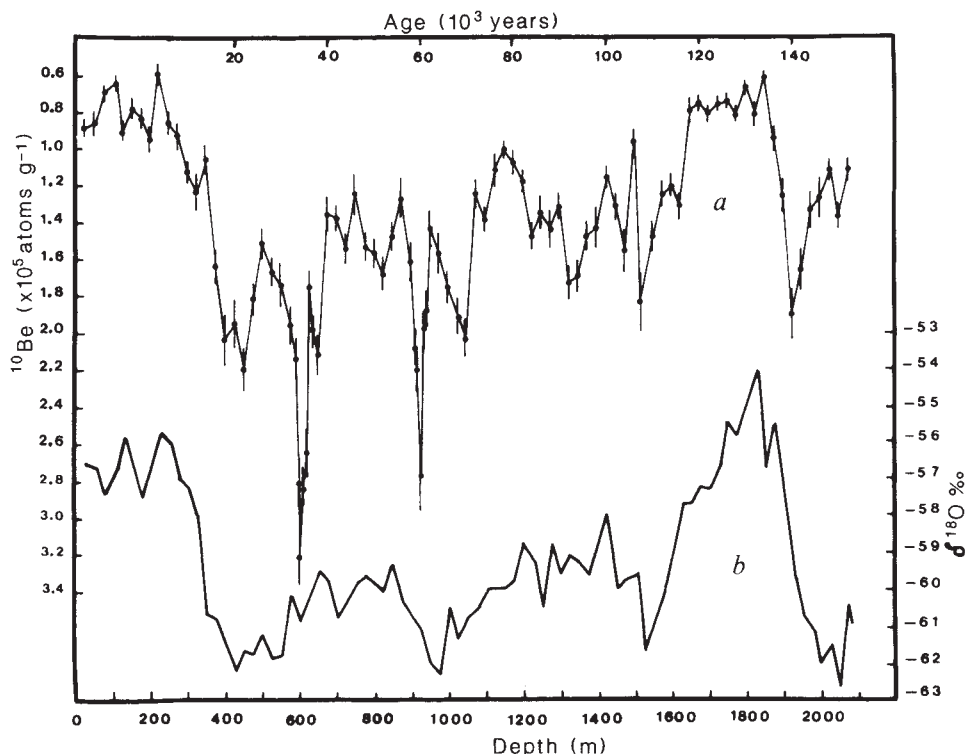
* Correct to ± 2 m.

† Estimated from chronology in ref. 3.

variations, or short-term solar modulation variations. For levels that were not subdivided, we have always tried to take samples corresponding to at least 10 yr of precipitation, in order to minimize the above effects. We thus believe that in general our measurements are reasonably representative of the ^{10}Be concentration at the levels sampled, although there may be occasional exceptions.

Our results, including those previously published², are shown in Fig. 1. For levels for which we have multiple measurements, we have plotted the weighted average value. The error bars indicate only estimated measurement uncertainties. As before² we have plotted the ^{10}Be on an inverted scale to facilitate comparison with the oxygen isotope curve³. The ^{10}Be profile in Fig. 1 in general confirms the trends noted in ref. 2: in particular

Fig. 1 *a*, The ^{10}Be concentration and *b*, $\delta^{18}\text{O}$ as a function of depth in Vostok ice core. Note inverted scale for ^{10}Be to facilitate comparison between the two curves. Small corrections ($<7\%$) for ^{10}Be decay have been made using the timescale (top abscissa) adopted by Lorius *et al.*³.



note the good correlation with the oxygen isotope curve, and the factor of ~ 2 larger ^{10}Be concentration during the glacial period compared to the Holocene and previous interglacial. In addition the greater detail of the present profile allows us to observe more subtle effects. For example, the ^{10}Be concentration appears to increase quite steadily from the previous interglacial until the glacial maximum at $\sim 20,000$ yr BP, followed by a rather rapid decrease to Holocene levels. Similar 'saw-tooth' structure is often observed in other climatological curves for this period⁴, although it is not particularly obvious in the published Vostok oxygen isotope record³. Also, there appear to be some differences in detail between the ^{10}Be and $\delta^{18}\text{O}$ curves for the period just before the previous interglacial, with the maximum in the ^{10}Be record being apparently much narrower than that of the $\delta^{18}\text{O}$. We will not comment further on these points here.

The most dramatic difference between the ^{10}Be and $\delta^{18}\text{O}$ curves are the two fairly well defined peaks in the ^{10}Be concentration, at ~ 925 m and ~ 600 m. According to the timescale adopted by Lorius *et al.*³, these levels correspond to times of $\sim 60,000$ and $\sim 35,000$ yr BP, respectively. Each of the peaks is estimated to last $\sim 1,000$ – $2,000$ years, depending on the assumed precipitation rate and how one chooses to define the peak shapes.

Once we discovered the ^{10}Be peak in the Vostok core at 600 m, we decided to measure ^{10}Be in samples from the same time period in the Dome C core, as estimated from a comparison of continuous $\delta^{18}\text{O}$ profiles in the two cores (the second peak at 925 m in the Vostok core is beyond the range covered by the Dome C core). A comparison of the Vostok and Dome C profiles is shown in Fig. 2. It is clear that both the relative amplitude and widths of the ^{10}Be peaks are very similar in the two cores. We believe this good correlation excludes experimental artefacts, or local geographical effects, as the cause of the observed ^{10}Be concentration increases. However, the fact that both these sites are on the Antarctic plateau, and subject to similar meteorological conditions, means that we cannot exclude regional climatic effects as a possible common cause of the peaks, as discussed below.

The first question we must address is whether the observed ^{10}Be concentration peaks are actually due to increased ^{10}Be deposition, or whether they are simply more dramatic examples of the precipitation variation effect mentioned earlier as being

the probable cause of the glacial-interglacial concentration differences. There are two arguments against this explanation. The first is that, according to our working hypothesis, which appears to be consistent with all the data collected so far^{1–3} the precipitation rate at these two sites seems to be controlled mainly by the water vapour pressure, and thus temperature, above the tropospheric inversion layer. The $\delta^{18}\text{O}$ data show no evidence for any major change in this temperature for the periods concerned. Second, and perhaps more convincing, any decrease in precipitation rate would be expected to show up also in increased concentrations of other aerosol components, such as dust and trace elements, for these periods. There is, however, no evidence for particular enhancements of these other components at these depths^{5,6}. Although both aluminum⁵ and microparticles (J. R. Petit, personal communication) show a concentration maximum at ~ 900 – $1,000$ m in the Vostok core, the width of the peak is considerably larger than that of the ^{10}Be . We thus conclude that the ^{10}Be peaks probably do reflect real increases in ^{10}Be deposition.

We have previously discussed two other potential causes for increased ^{10}Be concentration in polar ice¹. One of these is increased ^{10}Be deposition due to changes in atmospheric circulation. We estimate that only a rather small fraction of stratospherically produced ^{10}Be is presently deposited on the Antarctic plateau⁷, it is quite easy to imagine changes in this fraction giving rise to peaks of the type observed. One might again cite the absence of any correlation with other trace components in the ice as an argument against such an explanation. However, the source of these other components is quite different from ^{10}Be , so such an argument cannot be conclusive. Thus, although there is no evidence for changing atmospheric circulation as the source of the ^{10}Be peaks, ultimately the only way of excluding this explanation will be to find similar peaks in other reservoirs.

The third potential source of the ^{10}Be peaks, and the one we favour for the moment, is that they are due to increased production of ^{10}Be in the atmosphere. There are three possible causes for such production changes⁸: (1) variations in primary cosmic ray flux, (2) changes in solar modulation, (3) changes in geomagnetic field intensity.

Although there is still no general agreement on the source of galactic cosmic rays, it seems unlikely to us that variations of

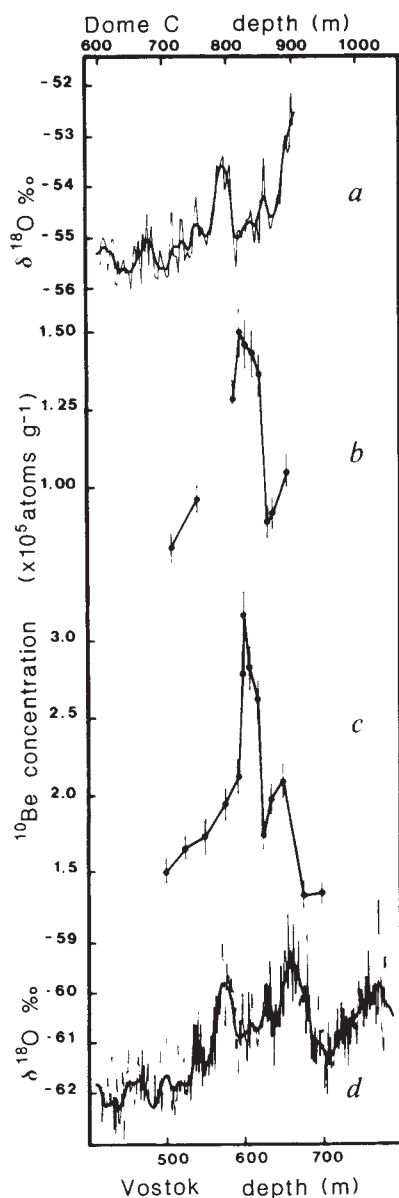


Fig. 2 a, $\delta^{18}\text{O}$ and b, ^{10}Be concentration in Dome C core; c, ^{10}Be concentration and d, $\delta^{18}\text{O}$ in Vostok core. For $\delta^{18}\text{O}$ the fine lines represent raw data, dark curves are fits with second order spline functions. The depths in the two cores have been correlated by the $\delta^{18}\text{O}$ curves, in particular the peaks at ~ 575 m and 650 m on the Vostok depth scale.

the magnitude necessary to produce the observed ^{10}Be peaks could take place on the timescales involved.

The reason we originally suggested² the possibility of changes in solar activity (through solar modulation) as the possible source of the observed ^{10}Be peak at 925 m (before we had any indication of its duration), is because changes in modulation are well documented at the ~ 11 yr timescale (solar sunspot cycle), and are also believed to be responsible for observed ^{14}C variations observed on timescales of ~ 100 – 200 years during the Holocene^{9,10}. Variations in ^{10}Be on similar timescales have also been attributed to this cause^{1,11,12}. If the ^{10}Be peaks presented here are due to changes in solar modulation, then they would represent evidence of reduced solar activity for much longer periods than previously documented. This could potentially have important implications for models of solar variability, and solar-terrestrial relationships.

The third possible source of ^{10}Be production variations are changes in the geomagnetic field intensity. It is widely believed

that such intensity changes were responsible for the long-term changes in the $^{14}\text{C}/^{12}\text{C}$ ratio in the atmosphere during the past $\sim 8,000$ years¹³. However, Beer *et al.*¹² have reported the absence of any corresponding variation in ^{10}Be in Greenland ice for the same time period. In order to increase global ^{10}Be production by a factor of ~ 2 , it would be necessary for the geomagnetic field intensity to fall to $<20\%$ of its present value¹⁴. Reductions of this magnitude are known to occur during geomagnetic reversals. We have in fact reported increased ^{10}Be concentrations in a marine sediment core at the time of the most recent of these reversals 700,000 years ago¹⁵. In addition to well recognized global reversals, there are numerous reports in the literature of geomagnetic 'excursions', periods during which the apparent dipole direction and/or intensity at a given location fluctuate strongly from their 'normal' values for periods of a few hundred to a few thousand years¹⁶. It should be emphasized, however, that there is no general agreement that any of these recorded excursions actually represent global changes in the geomagnetic field^{16,17}. In fact we have pointed out that ^{10}Be could be an important probe for testing the global nature of excursions¹⁵ (no significant changes in ^{10}Be production would be expected for only local or regional geomagnetic variations). Interestingly, several of the reported excursions (Laschamp¹⁸, Mono Lake^{19,20}, Lake Mungo²¹, Shibutami²²) have estimated ages in the range of 25,000–45,000 yr BP. It is thus natural to speculate whether the ^{10}Be peak observed here at $\sim 35,000$ yr BP could be caused by low fields associated with one or more of these excursions. (There is as yet no convincing evidence that any of these reported excursions are related^{16,17}.) There is at least one possible argument against such an explanation, and this is that the influence of the geomagnetic field on ^{10}Be production is strongest in the equatorial region, and essentially zero near the poles. Indeed, we have noted⁷ that this might be the reason that Beer *et al.*¹² did not observe, at Camp Century Greenland, the variation of ^{10}Be expected for a dipole change sufficient to give the observed long term variation of ^{14}C in the atmosphere during the Holocene. The ^{10}Be peaks observed here are larger than the calculated production increase in the polar regions, even for a dipole intensity of zero¹⁴. Thus, if geomagnetic variations are the source of the ^{10}Be peaks, it would imply that, at least during those periods, there was significant transport of lower latitude aerosols to the Antarctic plateau.

It is clear that considerably more work will be necessary to understand fully the implications of the periods of high ^{10}Be deposition presented here. Perhaps most urgent will be the research to try to find evidence of these periods in other geological reservoirs. It would be interesting, for example, to try to find similar ^{10}Be peaks at the appropriate depths in deep Arctic ice cores. High concentrations of ^{10}Be have been reported in late glacial ice at Dye 3, Greenland, but these have been correlated with low $\delta^{18}\text{O}$ values, and thus attributed as mostly resulting from low precipitation rates²³, although the authors did allow for possible contributions due to production or atmospheric circulation effects. We are also planning to investigate various lower-latitude high-deposition-rate marine and lacustrine sediment cores.

Regardless of the eventual source of these ^{10}Be variations, one potential application already seems assured, and that is their use as stratigraphic markers. At the very least, as shown by the evidence from the two cores studied here, these ^{10}Be events should permit correlation of deep ice over a substantial fraction of Antarctica. This could be important for locations for which meteorological or geographical conditions make correlation of oxygen isotope records ambiguous. If the ^{10}Be increases really do reflect production changes, then their presence in other reservoirs would permit a chronological connection between polar and lower latitude climatological records, as well as between the various lower latitude records themselves.

It is also obvious that any inferred change in production of ^{10}Be would have important implications regarding application

of other cosmogenic isotopes, such as ^{14}C .

Finally we note that, since the widths of the two ^{10}Be peaks observed here is comparable to our sampling interval ($\sim 25\text{ m}$) over most of the Vostok core, it is not impossible that there are additional peaks of similar or shorter duration, which we have missed. We thus hope to have the opportunity in the future of measuring an even higher resolution ^{10}Be profile in this core.

We thank J. Lestringuez for help in sample preparation and analysis. We also thank D. Deboffe, J. R. Petit and K. J. Blinov for logistical and technical help. Sample procurement was supported by the Soviet Antarctic Expeditions, Expéditions Polaires Françaises and the PNEDC (Programme National d'Études de la Dynamique du Climat). Tandetron operation is supported by the CNRS, CEA and IN2P3.

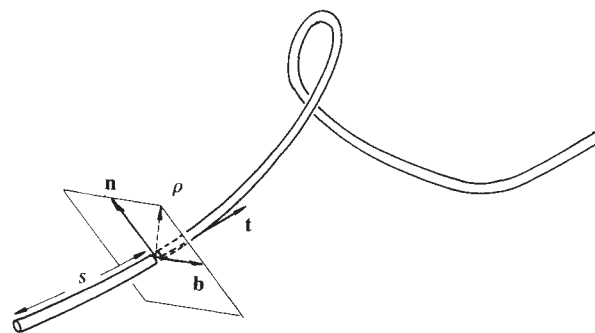


Fig. 1 Geometry of helically coiled optical fibre.

Received 2 December 1986; accepted 9 January 1987.

1. Raisbeck, G. M. *et al.* *Nature* **292**, 825–826 (1981).
2. Yiu, F., Raisbeck, G. M., Bourles, D., Lorus, C. & Barkov, N. I. *Nature* **316**, 616–617 (1985).
3. Lorus, C. *et al.* *Nature* **316**, 591–596 (1985).
4. Broecker, W. S. & Van Donk, J. *Rev. Geophys. Space Phys.* **8**, 169–198 (1970).
5. De Angelis, M., Barkov, N. I. & Petrov, V. N. *Nature* **325**, 318–321 (1987).
6. Petit, J. R., Briat, M. & Royer, A. *Nature* **293**, 384–391 (1981).
7. Raisbeck, G. M. & Yiu, F. *Ann. Glaciol.* **7**, 138–140 (1985).
8. Raisbeck, G. M. & Yiu, F. *Nucl. Instrum. Meth.* **B5**, 91–99 (1984).
9. Stuiver, M. & Quay, P. D. *Science* **207**, 11–19 (1980).
10. Neftel, A., Oeschger, H. & Suess, H. E. *Earth planet. Sci. Lett.* **56**, 127–147 (1981).
11. Beer, J. *et al.* *Radiocarbon* **25**, 269–278 (1983).
12. Beer, J. *et al.* *Nucl. Instrum. Meth.* **B5**, 380–384 (1984).
13. Damon, P. E., Lerman, J. C. & Long, A. A. *Rev. Earth planet Sci.* **6**, 457–494 (1978).
14. O'Brien, K. J. *J. geophys. Res.* **84**, 423–431 (1979).
15. Raisbeck, G. M., Yiu, F., Bourles, D. & Kent, D. V. *Nature* **315**, 315–317 (1985).
16. Verosub, K. L. & Banerjee, S. K. *Rev. Geophys. Space Phys.* **15**, 145–155 (1977).
17. Verosub, K. L. *Phil. Trans. R. Soc. A306*, 161–168 (1982).
18. Gillot, P. Y. *et al.* *Earth planet. Sci. Lett.* **42**, 444–450 (1979).
19. Liddicoat, J. C. & Coe, R. S. *J. geophys. Res.* **84**, 261–271 (1979).
20. Negrini, R. M., Davis, J. O. & Verosub, K. L. *Geology* **12**, 643–646 (1984).
21. Barbetti, M. F. & McElhinny, M. W. *Phil. Trans. R. Soc.* **281**, 515–542 (1976).
22. Tanaka, H. & Tachibana, K. *J. Geomagn. Geoelect.* **33**, 287–292 (1981).
23. Beer, J. *et al.* *Ann. Glaciol.* **5**, 16–17 (1984).

Interpreting the anholonomy of coiled light

M. V. Berry

H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK

Circular birefringence of purely geometric origin was recently predicted¹ and observed² in helically coiled monomode optical fibres, and widely reported^{3–5} as a successful application to photons of a general theory^{6,7} for phase shifts in adiabatically transported quantum states. However, earlier similar observations^{8–10} had been interpreted not by quantum mechanics but simply as a classical anholonomy, namely parallel transport of the polarization¹¹. Indeed, because the magnitude of the effect is independent of the wavelength of the light as well as Planck's constant, it might seem that 'classical' here means that not only quantum but also wave effects can be neglected. Here, I argue that these experiments, and their discrete analogues, are most appropriately described at the level of classical electromagnetism; the parallel transport law can then be derived (rather than assumed^{8–11}) and nonadiabatic polarization changes calculated.

In the quantum description¹, photons in right- or left-circularly polarized light are assumed to be in the eigenstate of positive or negative helicity defined by the local tangent vector $\mathbf{t}(s)$ of the fibre (Fig. 1). Because the input and output ends of the fibre are parallel, the eigenstate is transported round a closed loop in $\mathbf{t}$ space and thereby acquires the geometrical phase shift appropriate to spin one, namely⁶ minus or plus the solid angle Ω subtended by the loop at the origin of $\mathbf{t}$ space. These opposite phase shifts $\mp\Omega$ imply¹ that the direction of linear polarization will be rotated by Ω , and it is this gyrotropy that was observed^{2,8–10}.

Underlying this successful prediction there are, however, several uncertainties. In the absence of any obvious governing Hamiltonian it is not clear why the phase continuation rule of the general theory⁶ is applicable. Even if it is, the fact that the two helicity states are degenerate suggests that the application could equally be made to any superposition of them, yielding different results. (Relativistic arguments¹² do yield the phase continuation rule, and also uniquely select the helicity states, but are inapplicable within fibres.) Moreover, it is hard to see how this type of quantum description can provide estimates of the probability that the coiling will produce nonadiabatic transitions to the opposite helicity. In any case, the high photon flux in all the experiments so far carried out makes a quantum description unnecessary.

At the level of geometrical optics (shortwave limit) it is known^{13,14} that in a medium with smoothly-varying refractive index μ the electromagnetic field vectors are parallel-transported along a light ray. This result should, however, not be invoked to explain experiments with monomode fibres, because their fields cannot legitimately be described by ray optics.

In classical electromagnetism the field is governed by Maxwell's equations, with μ depending only on perpendicular distance ρ from the fibre axis. Using the ideas of coupled local modes and the weak-guidance approximation¹⁵, we can write the transverse electric field $\mathbf{E}$ at position s along the fibre as a superposition of fields linearly polarized along the fibre normal $\mathbf{n}(s)$ and binormal $\mathbf{b}(s)$ ¹⁶; this is an adiabatic approximation (with s playing the role that time does in quantum mechanics), valid for gently coiled fibres. Thus,

$$\mathbf{E}(\rho, s) = \exp\{i\beta s\}f(\rho)[c_1(s)\mathbf{n}(s) + c_2(s)\mathbf{b}(s)] \quad (1)$$

where β and $f(\rho)$ are, respectively, the propagation constant and modal amplitude appropriate to the straight fibre. Substitution into Maxwell's equation gives, after some analysis (to be published elsewhere), the following effective Hamiltonian evolution equation for the polarization coefficients c_1 and c_2 in terms of the fibre curvature κ and torsion τ :

$$i \frac{\partial}{\partial s} \begin{pmatrix} c_1(s) \\ c_2(s) \end{pmatrix} = \begin{pmatrix} \kappa^2(s)/2\beta & i\tau(s) \\ -i\tau(s) & 0 \end{pmatrix} \begin{pmatrix} c_1(s) \\ c_2(s) \end{pmatrix} \quad (2)$$

This simplest possible theory neglects radiative leaking, coupling to reflected modes and all elasto-optic effects.

Now $\beta \approx 2\pi/\lambda$ where λ is the light wavelength in the fibre cladding, and κ^{-1} and τ^{-1} are both comparable with the fibre bend distances, so in lowest approximation the term $\kappa^2/2\beta$ is negligible and the torsion terms dominate. Then, for initial linear polarization in direction α (relative to $\mathbf{n}$), (2) gives

$$c_2(s)/c_1(s) = \tan\left(\alpha - \int_{-\infty}^s ds'(\tau(s'))\right) \quad (3)$$

This is precisely the parallel transport law for $\mathbf{E}$, giving, after one complete helical turn, a polarization rotation equal to the

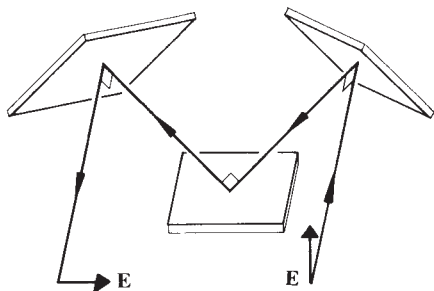


Fig. 2 Antiadiabatic rotation of polarization E by three reflections in metal mirrors.

solid angle¹⁰

$$\Omega = 2\pi - \int_{-\infty}^{\infty} ds' \tau(s')$$

Moreover, (2) resolves the degeneracy in favour of the helicity states: local eigenmodes (again neglecting $\kappa^2/2\beta$) are circularly polarized, that is, $c_2/c_1 = \pm i$, and acquire opposite phase shifts $\mp\Omega$ round the turn.

Nonadiabatic transitions are induced by the curvature term $\kappa^2/2\beta$. In lowest order this gives the probability for a change from (say) $+$ to $-$, as

$$P_{\mp} = \left| \int_{-\infty}^{\infty} ds \kappa^2(s) \exp \left\{ 2i \int_{-\infty}^s ds' \tau(s') \right\} \right|^2 / 16\beta^2 \quad (4)$$

For a helix uniformly wound on a cylinder of radius R so as to produce phase shifts $\mp\Omega$, this can be written as

$$P_{+-} = [\Omega/2\pi(2 - \Omega/2\pi)]^3 \sin^2 \Omega / (1 - \Omega/2\pi)^2 / 16\beta^2 R^2 \quad (5)$$

Even for $R \sim 1$ mm this never exceeds 10^{-8} . For planar bends ($\tau=0$) there is no polarization rotation, and the eigenmodes of (2) are linearly polarized with a tiny bend-induced shift $\kappa^2/2\beta$ in the local propagation constant of the mode polarized along $\mathbf{n}$.

It has been suggested (J. N. Ross, personal communication and ref. 17) that the coiling of $\mathbf{t}$ can be accomplished by (at least three) discrete reflections, the resulting sudden changes in $\mathbf{t}$ being simulations of adiabatic change. But ideal mirrors (infinite conductivity) do not conserve helicity; they reverse it, and with this 'antiadiabaticity' the solid angle Ω must be accumulated with $\mathbf{t}$ replaced by $-\mathbf{t}$ on alternate segments of the light path¹⁷. If the light beam is reversed after three successive reflections, each changing its direction by 90° (Fig. 2), the predicted polarization rotation (see also ref. 18, pages 84–86) is also 90° . Real metal mirrors (finite conductivity) do not quite reverse helicity: the 'nonadiabatic' probability for preserving the original helicity is

$$P_{++} \approx 1/(2|\mu|^2) \quad (6)$$

which for silver with $\mu = 0.2 + 3.44i$ is 0.042.

The simulation of true adiabatic change with discrete reflections can be achieved only by total internal reflection in a

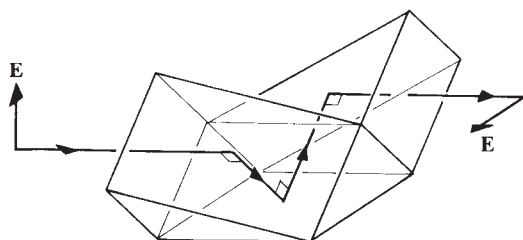


Fig. 3 Adiabatic rotation of polarization E by three internal reflections at the critical angle.

dielectric with index μ at the critical incidence angle $i = i_c = \sin^{-1}(\mu^{-1})$ (this follows from Fresnel's formulae¹³). For glass, $i_c \approx 45^\circ$, so that rotation of 90° can be accomplished by successive reflections in three 45° right prisms arranged as in Fig. 3. However, i_c is not precisely 45° (because $\mu \neq \sqrt{2}$), and this will cause nonadiabatic helicity switching, with probability

$$P_{+-} = (\mu^2 \sin^2 i - 1) / (\mu^2 \tan^2 i - 1) \quad (\sin i > \mu^{-1}) \quad (7)$$

For $i = 45^\circ$ and glass with $\mu = 1.5$, $P_{+-} = 0.1$.

The arrangements in Figs 2 and 3 form the basis of robust, simple and (nonadiabatic conditions notwithstanding) convincing lecture demonstrations of the anholonomic transport of polarization.

I thank Dr J. H. Hannay and Professor R. J. McCraw for helpful discussions, and to Professor M. Kitano for sending me his paper¹⁷ before publication.

Received 2 December 1986; accepted 3 February 1987.

- Chiao, R. Y. & Wu, Y. S. *Phys. Rev. Lett.* **57**, 933–936 (1986).
- Tomita, A. & Chiao, R. Y. *Phys. Rev. Lett.* **57**, 937–940 (1986).
- Maddox, J. *Nature* **323**, 199 (1986).
- Robinson, A. L. *Science* **234**, 424–426 (1986).
- MacCallum, M. *New Scientist* **1536**, 22 (1986).
- Berry, M. V. *Proc. R. Soc. A392*, 45–57 (1984).
- Simon, B. *Phys. Rev. Lett.* **51**, 2167–2170 (1983).
- Ross, J. N. *Optical Quantum Electron.* **16**, 455–461 (1984).
- Varnham, M. P., Birch, R. D. & Payne, D. N. *Tech. Dig. Int. Conf. Integrated Optics Optical Fiber Comms. Eur. Conf. Opt. Comms.* 135–138 (Istituto internazionale delle Comunicazioni, Genova, 1985).
- Varnham, M. P., Birch, R. D., Payne, D. N. & Love, J. D. *Tech. Dig. Conf. Opt. Fiber Comms. Atlanta* 68–71 (Optical Society of America, 1986).
- Haldane, F. D. M. *Optics Lett.* **11**, 730–732 (1986).
- Bialynicki-Birula, I. & Bialynicki-Birula, Z. *Phys. Rev. D* (in the press).
- Born, M. & Wolf, E. *Principles of Optics* (Pergamon, London, 1959).
- Kline, M. & Kay, I. M. *Electromagnetic Theory and Geometrical Optics* (Interscience, New York, 1965).
- Snyder, A. W. & Love, J. D. *Optical Waveguide Theory* (Chapman and Hall, London, 1983).
- Weatherburn, C. E. *Elementary Vector Analysis* (Bell, London, 1955).
- Kitano, M., Yabuzaki, T. & Ogawa, T. *Phys. Rev. Lett.* **58**, 523 (1987).
- Chow, W. W., Gea-Banacloche, J., Pedrotti, L. M., Sanders, V. E., Schleich, W. & Scully, M. O. *Rev. mod. Phys.* **57**, 61–104 (1985).

Evidence of recent lead pollution in deep north-east Atlantic sediments

Alain Veron*, Claude E. Lambert*, Ann Isley*, Philippe Linet† & Francis Grousset‡

* Centre des Faibles Radioactivités, laboratoire mixte CNRS/CEA, Domaine du CNRS, 91190 Gif-sur-Yvette, France

† Service Central d'Analyses, CNRS, BP22, 69390 Vernaison, France

‡ Département de Géologie et Océanographie, Université de Bordeaux 1, Avenue des Facultés, 33405 Talence, France

Today, the global lead emissions into the atmosphere from anthropogenic sources are about 200 times higher than in the past¹ and dominate other lead pathways to the ocean. Exhaust from leaded gasoline is the major source of pollutant lead in the atmosphere², and subsequently in the marine environment due to atmospheric deposition by wet and dry removal processes^{3,4}. This pollution has already perturbed the biogeochemical cycle of lead in the open ocean^{2,3,5–7} especially in the north Atlantic. In surface waters, where lead is introduced mainly in the dissolved state, its residence time is relatively short (less than 5 years), as it is rapidly scavenged by, and transported with, particulate matter^{5,7}. The mean residence time of lead in the ocean is estimated to be ~ 100 yr^{5,8}. One can therefore expect to see some pollutant lead in the deep ocean surficial sediments. Here, we report lead concentrations in two short cores from the north-east Atlantic. Lead concentrations in surficial sediments (21.0 and 15.0 p.p.m. in the top cm) are higher than in the uppermost 10 cm of the cores (6.0 and 2.8 p.p.m.), and are likely to be enhanced from anthropogenic sources. The quantity

of lead stored in these surficial sediments (5.7 and $2.5 \mu\text{g cm}^{-2}$) is of the same order as the amount of pollutant lead present in the dissolved state in the water column.

In this work, we try to establish the degree of lead pollution in two cores of eastern Atlantic sediments sampled during the Fluxatlante cruise in spring 1985 on the RV *Le Suroit*. The sediments were characterized by CaCO_3 content, accumulation rate from ^{14}C records, planktonic *foraminifera* biostratigraphy and ^{210}Pb excess ($^{210}\text{Pb}_{\text{xs}}$) in surficial sediments; this $^{210}\text{Pb}_{\text{xs}}$ can originate from atmospheric input or production in the water column⁹. The two stations were selected for their different environmental conditions. Station 7 ($46^\circ 22' \text{N}$, $12^\circ 23' \text{W}$), located on the Armorican Seamount at $3,650 \text{ m}$ depth, is several hundred metres above the abyssal plain and about 250 miles from the nearest continental landmass. The sediment is largely biogenous. Station 8 ($47^\circ 30' \text{N}$, $08^\circ 31' \text{W}$) is located on the Meriadzeck Plateau at $2,200 \text{ m}$ depth, on the continental slope, about 150 miles from the French coast, and is influenced much more by down-slope movement of silicate material.

The "Interface" core system (CNRS, Bordeaux) was used to recover sediment samples. The corer allows simultaneous recovery of three short-barrel (length: 50 cm) cores. The inner core barrel is composed of rigid polyvinyl chloride to minimize metallic contamination of the sediments. Examination of the cores upon recovery aboard ship demonstrated the presence of fragile surface dwellers, small-scale (cm) surface structures and particulate fluff, indicating good recovery of the sediment surface. The cores were immediately frozen aboard ship, and later sectioned in $3\text{- to }10\text{-mm}$ slices using a circular stainless steel saw. The sediment in direct contact with the saw was discarded ($\sim 1 \text{ mm}$). The sediment was dried for 48 h at 60°C and then crushed using an agate mortar and pestle. The crushed sediment was split into two homogeneous samples and transferred to pre-cleaned plastic Petri dishes until acid dissolution and analysis were carried out under clean conditions. Lead was determined by graphite furnace atomic absorption spectrometry and potentiometric stripping voltametry. The two techniques showed good agreement. The precision of these techniques was $\pm 10\%$; the blanks showed lead concentrations less than 10% of the lowest lead concentration in the samples. The accuracy was checked by analysing lead in NBS standards. Spectrometric measurements of naturally-occurring low-energy γ -rays were made using a high precision Ge detector to determine the activities of ^{210}Pb and ^{226}Ra ¹⁰.

The distribution of total Pb at both stations is shown in Fig. 1a. The two stations show an enrichment of Pb in surficial sediment (first centimetre). Below this, Pb concentrations in each core are low and show little variation with depth; these Pb background concentrations are $2.8 \pm 0.5 \text{ p.p.m.}$ at station 7 and $6.0 \pm 0.9 \text{ p.p.m.}$ at station 8. $^{210}\text{Pb}_{\text{xs}}$ could only be detected in the first centimetre of both cores indicating the penetration depth due to bioturbation over the last century (about five half-lives of ^{210}Pb).

The main difference between the cores appears in the sedimentation rate: 2.2 and 4.5 cm kyr^{-1} for stations 7 and 8. Taking into account a bulk density of 1.45 and 1.54 g cm^{-3} in surface cores at stations 7 and 8, and a calcium carbonate content averaging 87% at station 7 and 63% at station 8, we determined for stations 7 and 8 an average flux of 180 and $1450 \mu\text{g cm}^{-2} \text{ yr}^{-1}$ respectively of silicate material and of $1,800$ and $2,470 \mu\text{g cm}^{-2} \text{ yr}^{-1}$ of CaCO_3 .

First, we consider the background concentration of lead and the reason for a difference of a factor of two in the lead background values between the two cores. Using regressions on each core, we calculated for subsurface samples a theoretical Pb content of $15 \pm 3 \text{ p.p.m.}$ in the silicate fraction and $0.8 \pm 0.6 \text{ p.p.m.}$ in the carbonate fraction. These values are consistent with a world average estimate of 12 p.p.m. in the silicate fraction¹¹ and 0.1 to 1 p.p.m. in the carbonate fraction¹²⁻¹⁴. For subsurface samples, the concentrations of Pb fit satisfactorily

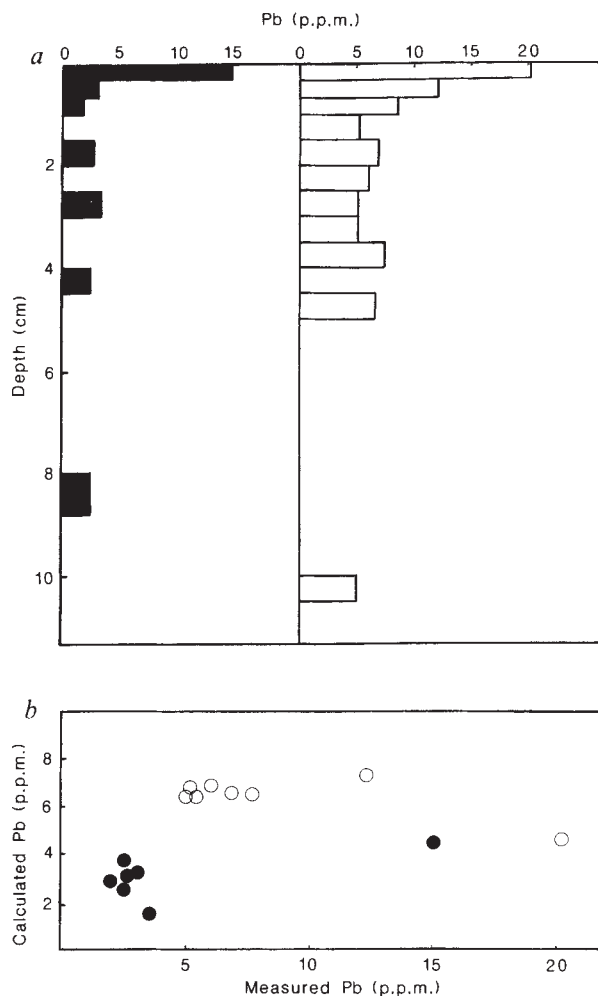


Fig. 1 a, Distribution of total Pb ($\mu\text{g g}^{-1}$ of dry sediment) at station 7 (■) and at station 8 (□). b, The calculated concentrations of Pb (using 0.8 p.p.m. Pb in carbonate fraction and 15 p.p.m. Pb in silicate fraction) is plotted against the measured values. The two groups of data represent background Pb concentrations at station 7 (●) and at station 8 (○) respectively. These data can be explained by the relative concentrations of CaCO_3 in the two cores. The isolated data represent surface Pb concentrations; they cannot be explained by the calculated data and therefore Pb must have another origin.

those predicted for a mixture of carbonate and silicate material with theoretical Pb contents of 15 and 0.8 p.p.m. (Fig. 1b). The surficial samples, however, show a significant excess of Pb above the values predicted on this basis (Fig. 1b). Considering these results, we can verify that the difference in background lead concentrations between the two stations can indeed be explained by different contributions of carbonate and silicate material in each core (Fig. 1b). The background level of lead, as defined here, does not necessarily represent a 'natural' level: indeed, lead pollution has existed for more than $2,000$ years. Therefore the background level we refer to here may include some old pollutant lead.

We consider that the Pb concentrations in subsurface sediments correspond to a 'background' level, and we assume that there is no important migration of lead in pore waters: this latter assumption is supported by the fact that isotope ratios of lead in coastal surface sediments differ from subsurface ratios and are close to that observed in the atmospheric input².

For our two cores we observe a low organic carbon content (4 mg g^{-1} of dry sediment) in the subsurface, characteristic of oxic conditions^{15,16}. This observation is confirmed by low concentrations of particulate manganese measured at the interface

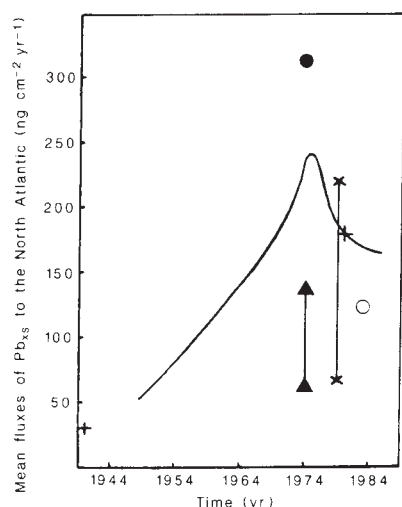


Fig. 2 Mean atmospheric fluxes of pollutant lead (Pb_{xs}) to the north Atlantic. Symbol —, from gasoline exhaust data, see calculation in text, data from Octel, France, and Comité professionnel du Pétrole, Paris, France. Symbol +, (ref. 4); Δ , (ref. 29); $\times$, (ref. 30); $\circ$, (ref. 22); $\bullet$, (ref. 24).

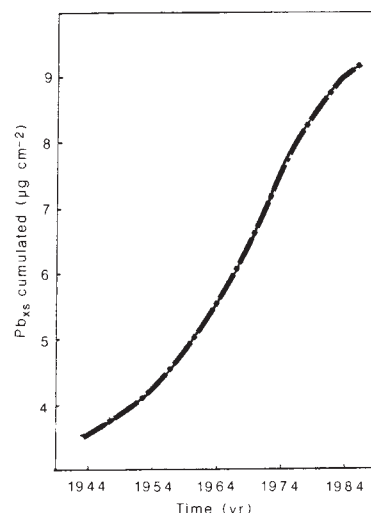


Fig. 3 Average inventory of the input of anthropogenic Pb over the north Atlantic, including Pb accumulated before 1940 ($3 \mu g cm^{-2}$, see text) and after 1940 (from Fig. 2).

on a free-carbonate basis (1,000 and 3,160 p.p.m. at stations 7 and 8 respectively). Therefore there is probably little reduction of Mn and Fe oxides in these cores¹⁶ and the excess lead observed in surficial sediments is unlikely to be due to lead migration. Then the excess lead over background (Pb_{xs}) observed in the first centimetre (Fig. 1a) could be of recent anthropogenic origin. This is consistent with the fact that the penetration depth of this Pb_{xs} is the same as that of $^{210}Pb_{xs}$ due to bioturbation in the past hundred years.

Another argument for a recent input of anthropogenic lead in surficial sediments is based on existing sediment trap data, although most existing sediment trap material has not been suitable for lead measurements. However, one data set is reported in the Sargasso Sea at 3,200 m depth¹⁷. It has been shown that the total flux at 3,200 m responds rapidly to surface conditions and that little time elapses between the input of material in surface waters and the flux at great depths^{19–23}. The reported deep-sea flux for 1982 ($90 ng cm^{-2} yr^{-1}$) is in the range of the atmospheric Pb fluxes reported for adjacent waters^{22,23}, and one can infer that 60–80% of the anthropogenic yearly input in the Sargasso Sea is reaching the sediment.

The Pb_{xs} inventory, taking into account porosity measurements in each core, is $2.5 \mu g cm^{-2}$ at station 7 and $5.7 \mu g cm^{-2}$ at station 8. At station 7, the silicate flux corresponds to $180 \mu g cm^{-2} yr^{-1}$ of silicate material accumulation compared to an open ocean atmospheric input of $50–130 \mu g cm^{-2} yr^{-1}$ (refs 24, 25). One can therefore consider that most of the silicate material comes from the atmosphere and infer that most of the lead also comes from the atmosphere which is its main input source. Station 8 is on the slope, closer to the European coast, and subjected to more input from the upper slope, as evidenced by a silicate flux eight times more important at station 8 than at station 7. Geographical variations in the sediment record are to be expected, as it has been shown that Pb is more rapidly scavenged at the boundaries than in the open ocean³. How does this inventory of Pb_{xs} compare with the water column inventory and the atmospheric input over the Atlantic?

The dissolved lead of recent anthropogenic origin in the north Atlantic water column can be calculated as equal to Pb concentrations above $5 ng l^{-1}$ (estimated background concentration). According to the Schaule and Patterson Atlantic profile²⁶, it amounts to about $3 \mu g cm^{-2}$ above the background. This represents 50–100% of the inventory of our cores.

There is little available data on the atmospheric flux of Pb over the north Atlantic. We estimated an average flux for the

last 40 yr (Fig. 2) using the data on lead in gasoline (OCTEL) and assuming that two-thirds of the Pb from gasoline exhausts escapes into the atmosphere. The data obtained, from Europe, Canada and USA consumption, were adjusted to the Settle and Patterson Pb atmospheric fluxes for 1981⁴. The other published data from atmospheric lead concentration measurements or estimates are within a factor of two of the calculated average values based on Pb emission data. To calculate an average pre-war industrial lead input over the Atlantic, we used the average 1940 atmospheric deposition flux calculated by Flegal and Patterson⁷ and the slope from the data of Murozumi *et al.*²⁷ showing lead accumulation in Greenland ice from 1750 to 1940. This leads to an average inventory of $3 \mu g cm^{-2}$ 'old' industrial Pb before 1940.

Adding both the recent lead input (essentially from gasoline exhaust) and this older industrial input (Fig. 3), we estimate an average accumulated anthropogenic lead inventory of about $9 \mu g cm^{-2}$ over the north Atlantic. Considering the inventories of our two cores, ~30–60% of the average input is recorded in the sediment.

Therefore it appears that the two estimates of the north Atlantic atmospheric input from: (1) available data in the atmosphere and (2) Pb inventory in the water column and our results in the sediments are in good agreement. This strongly suggests that a large fraction of this pollutant lead is already accumulated in deep north-east Atlantic sediments. If confirmed, this conclusion should be taken into account when quantitatively modelling the future response of the lead cycling in the ocean to changes in the atmospheric input function. Such changes are already occurring due to the phasing out of leaded gasoline in the USA and have recently been recorded in surface marine waters²⁸.

We thank F. Dehaers for $CaCO_3$ and carbon determinations and R. C. Tian for providing ^{14}C accumulation rates. Helpful discussions with T. Jickells, P. Buat-Menard, S. Calvert and reviewers comments are acknowledged. This work was supported by funds from CEA, CNRS and ATP Chemical Oceanography and Global Fluxes (FMO).

Received 9 September 1986; accepted 9 January 1987.

1. Settle, M. D. & Patterson, C. C. *Science* **207**, 1167–1176 (1980).
2. Ng, A. & Patterson, C. C. *Geochim. cosmochim. Acta* **46**, 2307–2321 (1982).
3. Schaule, K. B. & Patterson, C. C. *Earth planet. Sci. Lett.* **54**, 97–116 (1981).
4. Settle, D. M. & Patterson, C. C. *J. geochem. Res.* **87**, 11, 8857–8869 (1982).

5. Flegel, A. R. & Patterson, C. C. *Earth planet. Sci. Lett.* **64**, 19–32 (1983).
6. Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* **26**, 263–308 (1962).
7. Burnett, M. W. & Patterson, C. C. in *Isotope Marine Chemistry* (eds Goldberg, E. D., Horibe, Y. & Saruhashi, K.) 413–438 (Uchida Rokakuho, Tokyo, 1980).
8. Bacon, P. M. thesis, Woods Hole Oceanogr. Inst. (1976).
9. Cochran, J. K. in *Uranium Series Disequilibrium: Applications to Environmental Problems* (eds Ivanovich, M. & Harmon, R. S.) 384–430 (Clarendon, Oxford, 1982).
10. Yokohama, Y. & Nguyen, H. V. in *Isotope Marine Chemistry* (eds Goldberg, E. D., Horibe, Y. & Saruhashi, K.) 259–289 (Uchida Rokakuho, Tokyo, 1980).
11. Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273–1285 (1964).
12. Marshall, R. R. & Patterson, C. C. in *Annual Report, Amer. Petroleum Inst. Project 53* (Calif. Inst. of Tech., Pasadena, 1955).
13. Shen, G. T. & Boyle, E. A. *Eos* **65**, 964 (1986).
14. Thompson, G. *Geochim. cosmochim. Acta* **36**, 1237–1253 (1972).
15. Froelich, P. N. *et al. Geochim. cosmochim. Acta* **43**, 1075–1090 (1979).
16. Bender, M. L. & Heggie, D. T. *Geochim. cosmochim. Acta* **48**, 977–986 (1984).
17. Jickells, T. D., Deuser, W. G. & Knap, A. H. *Deep Sea Res.* **31**, 1169–1178 (1984).
18. Deuser, W. G., Brewer, P. G., Jickells, T. D. & Commeau, R. F. *Science* **219**, 388–391 (1983).
19. Deuser, W. G., Ross, E. H. & Anderson, R. F. *Deep Sea Res.* **28**, 495–505 (1981).
20. Deuser, W. G. & Ross, E. H. *Nature* **283**, 364–365 (1980).
21. Bacon, M. P., Huh, L. A., Fleer, A. P. & Deuser, W. G. *Deep Sea Res.* **32**, 273–286 (1985).
22. Jickells, T. D., Knap, A. H. & Chuch, T. M. *J. geophys. Res.* **89**, 1423–1428 (1984).
23. Chuch, T. M. *et al. Atmos. Environ.* **18**, 2657–2664 (1984).
24. Buat-Menard, P. & Chesselet, R. *Earth planet. Sci. Lett.* **42**, 399–411 (1979).
25. Duce, R. A., Ray, B. J., Hoffman, G. L. & Walsh, P. R. *Geophys. Res. Lett.* **3**, 339–342 (1976).
26. Schaule, B. K. & Patterson, C. C. in *Trace Metals in Seawater* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) (NATO Conference series, Plenum Press, New York, 1983).
27. Murozumi, M., Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* **33**, 1247–1294 (1969).
28. Boyle, E. A., Chapnick, S. D., Shen, G. T. & Bacon, M. J. *geophys. Res.* **91**, 8573–8593 (1986).
29. Duce, R. A. *et al. in Marine Pollutant Transfer* (eds Windom, H. & Duce, R.) 77–119 (Lexington, 1976).
30. Chester, R., Sharples, E. J., Murphy, K., Saydam, A. C. & Sanders, G. S. *Mar. Chem.* **13**, 57–72 (1983).

Ultra-high resolution ^{29}Si MAS NMR spectra of highly siliceous zeolites

C. A. Fyfe*, J. H. O'Brien & H. Strobl

Guelph-Waterloo Centre for Graduate Work in Chemistry,
Guelph Campus, Department of Chemistry and Biochemistry,
University of Guelph, Guelph, Ontario N1G 2W1, Canada

By preparing the most highly crystalline samples possible and very carefully optimizing all the experimental variables, ultra-high resolution ^{29}Si magic-angle spinning nuclear magnetic resonance spectra have been obtained for completely siliceous zeolite ZSM-5. The spectra show a degree of resolution hitherto not thought possible for rigid crystalline materials, having natural linewidths without any spectral enhancement of ~5 Hz (~0.06 p.p.m.), an improvement of an order of magnitude, and exhibit clearly resolved resonances for 21 of the 24 postulated lattice sites. The resolution is largely maintained with variation of temperature and in the presence of organic sorbates, making it possible to identify clearly the numbers of lattice sites in the resulting new lattice structures.

In our recent investigations of zeolite structures by high-resolution solid state magic-angle spinning nuclear magnetic resonance (MAS NMR), we have found it particularly informative to study completely siliceous analogues formed by hydrothermal dealumination techniques^{1,2}. The complete removal of aluminium from the zeolite framework yields very narrow resonances in the ^{29}Si MAS NMR spectrum which are due to the crystallographically inequivalent silicon sites in the lattice structure. Taken together with X-ray diffraction data, the spectra provide a more complete description of zeolite structures, probing such subtle effects in some zeolites as lattice distortions and changes in the lattice structure induced by temperature and by the presence of organic molecules, as well as revealing ambiguities in the spectral determination of the Si/Al ratio of the low Si/Al ratio analogues. In the study of these highly siliceous materials, it is critical that the highest possible resolution be obtained for a particular sample. In this report, we discuss those factors which we have found to be most critical in this regard and present current results for zeolite ZSM-5 under various conditions for use as reference data by other

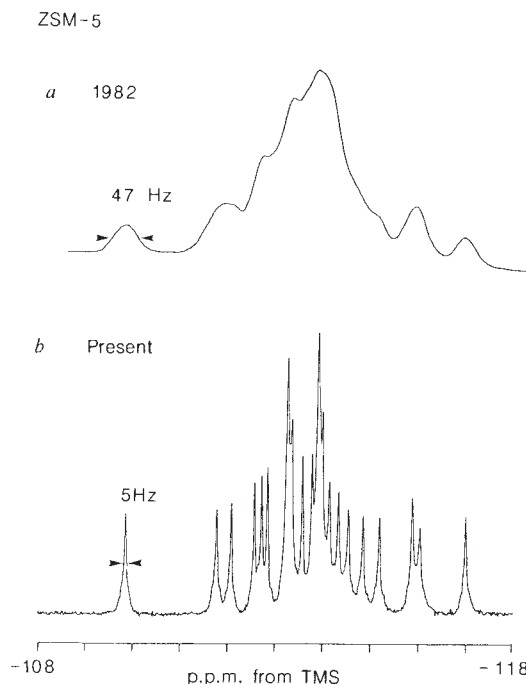


Fig. 1 *a*, ^{29}Si MAS NMR spectrum of highly siliceous ZSM-5 obtained at 79.5 MHz, 6,550 scans, 0 Hz linebroadening. Reproduced by permission from ref. 4. *b*, ^{29}Si MAS NMR spectrum obtained at 79.5 MHz, 1,633 scans, 0 Hz linebroadening (TMS, tetramethyl silane).

workers in this area. Spectra of comparable resolution may be obtained for other zeolite systems.

We have found it very beneficial to have as intimate a connection between the zeolite syntheses and the NMR spectroscopy as possible. After 'in house' preparation, which permits the deliberate synthesis of very highly crystalline, very high Si/Al ratio materials, the samples are prescreened by XRD measurements and only the very best selected for further treatment. The sample of ZSM-5 used in this present study was prepared by taking four separate solutions A–D consisting of: A, 10 g ammonium bifluoride in 20 ml water; B, 36 Ludox in 123 ml water; C, 6.5 g sodium hydroxide in 10 ml water and D, 10.6 g tetrapropylammonium bromide in 20 ml water. A and B were combined then C and finally D added and the resulting mixture (pH = 8.0–8.5) heated in a stainless steel bomb at 170 °C for 5 days. In the hydrothermal dealumination process, there will be optimum treatment times and temperatures which must be determined. The dealumination conditions used for the present sample were a single continuous treatment at 800 °C for nine days using air saturated with 10% aqueous HCl. The results obtained are for the most part reproducible.

The static magnetic field was shimmed on a liquid sample oriented at the magic angle using the ^{27}Al resonance of 1 mol l⁻¹ Al³⁺ (aq) solution. Homogeneity of ~10 Hz ($\Delta\nu_{1/2}$) could be easily obtained at 400 MHz over the 600 μl sample volume used. Shimming on the ^2H resonance of D₂O yields $\Delta\nu_{1/2} \approx 2$ –3 Hz. The resulting homogeneity of a spinning sample may be conveniently checked using a spinner filled with a cured polydimethylsiloxane elastomer (a variety of commercial caulking products of this material area readily available). For long accumulations, field stabilization using an external ^2H lock may be necessary.

Very-stable-sample spinning was obtained using a modified version of a spinner system by Wind (R. Wind, personal communication). The cylindrical nature of this design also facilitated shimming the magnetic field. Angle adjustment was made by optimizing the sideband intensities of the outer transitions of the ^{81}Br resonance of KBr mixed with the sample as described

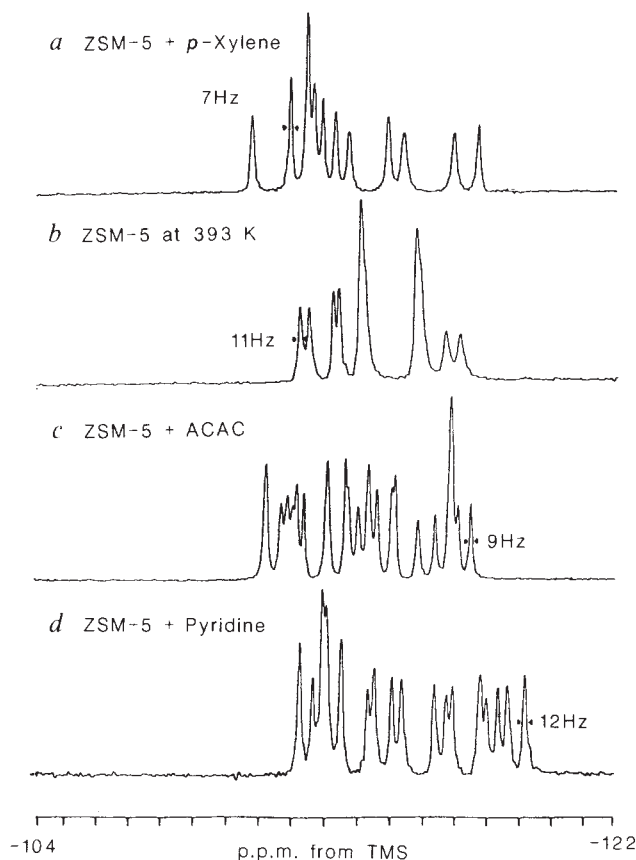


Fig. 2 ^{29}Si MAS NMR spectrum of highly siliceous ZSM-5 obtained at 79.5 MHz. *a*, At 295 K, after the addition of 40 μL of *p*-xylene per gram of zeolite, 500 scans, 0 Hz linebroadening. *b*, At 393 K, 1,225 scans, 0 Hz linebroadening. *c*, At 295 K, after the addition of 80 μL of acetylacetone per gram of zeolite, 4,185 scans, 0 Hz linebroadening. *d*, At 295 K, after the addition of 140 μL of pyridine per gram of zeolite, 1,122 scans, 0 Hz linebroadening.

by Frye and Maciel³. A lever system using a screw thread was used for fine adjustment of the spinning angle with one full rotation of the adjusting rod corresponding to ~ 1.7 degrees.

The effect of careful optimization of all of the factors discussed above is illustrated in the spectra of the highly siliceous ZSM-5 zeolite in Fig. 1. Figure 1*a* shows the spectrum presented in our original investigation of this system⁴. All of the resonances are due to $\text{Si}[4\text{Si}]$ environments and reflect the unique local geometries within the ZSM-5 unit cell. The resolution is good enough that the two outer resonances are well defined, allowing the deduction that the unit cell must contain 24 silicon atoms in the basic structural unit. However, only a maximum of nine peaks can be said to be actually observed, although the linewidths are ≈ 47 Hz (~ 0.6 p.p.m.) which is very narrow for a rigid solid. Similar spectra have subsequently been obtained by other workers⁵⁻⁷. Since then, we have gradually improved the spectral resolution in this system, confirming the relative intensities of the outer resonances in the spectrum⁸. By use of the techniques described above, these improvements have culminated in the spectrum shown in Fig. 1*b* where all of the experimental variables have been carefully optimized. The linewidths have now been reduced to ~ 5 Hz (0.06 p.p.m.), and provide a degree of resolution hitherto not thought possible for rigid crystalline materials. Twenty-one signals are now clearly observed which can be easily and unambiguously curve resolved. It should be emphasized that no resolution enhancement procedures have been applied to the data presented here. It is anticipated that even further gains in resolution with this sample

might be made at higher field strengths and additional increases obtained from the application of resolution enhancement techniques and from alternative data treatments to the fast Fourier transform (FFT) procedure. Experiments in these areas are currently in progress. The main point is that the ultimate resolution obtainable for these very crystalline lattice structures (< 0.1 p.p.m.) is very much greater than previously anticipated, making the MAS NMR technique a much more sensitive probe of the lattice structures than was thought to be the case.

The increased resolution is maintained both at elevated temperatures and in the presence of sorbed organic molecules. Thus, on adding *p*-xylene to the system a new spectrum is produced in which 11 lines are clearly resolved (Fig. 2*a*) whose relative intensities indicate 12 independent T sites in the unit cell in agreement with a symmetry change from monoclinic to orthorhombic. Previous spectra resolved only a few of the resonances⁹. In addition, the ultra-high resolution presently obtainable now establishes that the symmetry change is dependent on the organic substrate. There are 24 equivalent T sites in the new lattice structures formed when pyridine and acetylacetone are the sorbates (Fig. 2*c*, *d*).

Greatly improved resolution is also obtained at elevated temperatures although it is somewhat worse than at ambient. The limiting spectrum at 120 $^{\circ}\text{C}$ (393 K) (Fig. 2*b*) is similar in general form to those previously obtained above the phase transition^{1,2,10,22} but now some ten resonances are detectable whose relative intensities clearly indicate a twelve tetrahedron basic building unit, again compatible with orthorhombic or pseudo-orthorhombic symmetry, resolving an ambiguity in previous investigations¹¹. From spectra obtained at different temperatures, all twelve resonances can be identified.

The phase behaviour of the ZSM-5 lattice under the combined effects of temperature and sorbed organic molecules is exceedingly complex and the ultra-high resolution ^{29}Si MAS NMR spectra which we can now obtain reflect the lattice changes unambiguously and should be invaluable in probing these changes. We hope to improve the quality of the X-ray diffraction determined lattice structures of these systems by using synchrotron radiation and feel that a clear understanding of the phenomenon will result from the combined use of improved data from these two complementary techniques.

It should be possible to attain this degree of resolution for very highly siliceous zeolites in general and we have achieved this for several other systems to date. The present spectra are presented for use as benchmark references in studies of these systems by other workers in the area.

A further implication of these results is that it should be possible to obtain similar improvements in the resolution of the ^{13}C spectra of many crystalline organic materials although local susceptibilities of π -bonded carbons will still limit the linewidths of some resonances. The difference between the present MAS experiments and the conventional CP/MAS (CP, cross polarization) technique is that proton decoupling is not necessary, suggesting that the current limit of resolution in the CP/MAS experiment for very crystalline solids may be the efficiency of the proton dipolar decoupling. Experiments to investigate this possibility are currently in progress.

We acknowledge the financial assistance of NSERC (Canada) in the form of an operating grant (C.A.F.) and the technical assistance of Ms G. E. Barlow in the preparation of the sample.

Received 19 September; accepted 3 December 1986.

1. Fyfe, C. A. *et al.* in *New Directions in Chemical Analysis* 25-55 (Texas A&M Press, College Station, Texas, 1985).
2. Kokotailo, G. T. *et al.* *Pure Appl. Chemistry* **58**(10), 1367-1374 (1986).
3. Frye, J. S. & Maciel, G. E. *J. mag. Res.* **48**, 125-131 (1982).
4. Fyfe, C. A., Gobbi, G. C., Klinowski, J., Thomas, J. M. & Ramdas, S. *Nature* **296**, 533-536 (1982).
5. Nagy, J. P., Gabelica, Z., Deruane, E. G. & Jacobs, P. A. *Chem. Lett.* 2003-2006 (1982).
6. Smith, J. V. & Blackwell, C. S. *Nature* **303**, 223-225 (1983).
7. Boxhoorn, G., Kortbeek, A. G. T. G., Hays, G. R. & Alma, N. C. M. *Zolites* **4**, 15-21 (1984).

8. Fyfe, C. A., Gobbi, G. C. & Kennedy, G. J. *J. phys. Chem.* **88**, 3248–3253 (1984).
9. Fyfe, C. A., Kennedy, G. J., DeSchutter, C. T. & Kokotailo, G. T. *JCS Chem. Commun.* 541–542 (1984).
10. Fyfe, C. A., Kokotailo, G. T., Kennedy, G. J., Lyster, J. R. & Fleming, W. W. *JCS Chem. Commun.* 740–742 (1985).
11. Hay, D. G., Jaeger, H. & West, G. W. *J. phys. Chem.* **89**, 1070–1072 (1985).

Cerium isotope geochemistry of ocean island basalts

A. P. Dickin*

Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

The source of ocean island basalts is one of geology's major problems, and many different models have been proposed. One of these involves recycling of continental and oceanic crustal material into the mantle source of ocean island basalt magmas^{1,2}. Island arc basalts are an important link in the chain of evidence that supports this model, as they display isotope signatures in common with both ocean-floor sediment^{2,4} and ocean island basalts^{1,5}. But hafnium isotope data⁶ appear either to place severe restrictions on the quantity of sedimentary material recycled into the mantle, or to demand very thorough homogenization of this material during the subduction process. The lanthanum–cerium isotope method⁷ offers a prospect of applying similar constraints to the recycling of different sediment types into the mantle, based on cerium and neodymium isotope systematics. Here I report cerium isotope ratios for eight ocean island basalts. When plotted against published neodymium isotope data these form a 'mantle array' of compositions which indicates coherent behaviour of the two elements during mantle differentiation. This precludes the recycling into the deep mantle of ocean-floor sediments with a large net cerium anomaly.

Cerium undergoes fractionation in the marine environment relative to the other rare-earth elements, such that it undergoes relative depletion in deep ocean water and pelagic sediment, but enrichment in ocean-floor manganese nodules⁸. This fractionation is attributed to the removal of dissolved trivalent cerium by oxidation to a highly insoluble tetravalent state, not observed for other rare-earth elements.

Basaltic lavas from the Mariana Arc display negative cerium anomalies of ~20% relative to other light rare-earths⁹, suggesting that cerium-anomalous sediment is subducted into the mantle. If such material were carried down towards the lower mantle it might subsequently be sampled by ocean island basalt (OIB) magmatism.

Roden *et al.*¹⁰ reported a negative cerium anomaly in a tholeiitic basalt from the Koolau series volcanics of Hawaii, and similar anomalies have subsequently been found in East Molokai volcanics¹¹ and West Molokai volcanics (D. A. Clague and F. A. Frey, unpublished data). If these anomalies reflect ancient material subducted into the OIB source then the greater La/Ce ratios in these samples would have generated higher ¹³⁸Ce/¹³⁶Ce isotope ratios than in samples with normal rare-earth element profiles, due to the decay of ¹³⁸La to ¹³⁸Ce over geological time.

Lead isotope data yield an apparent age for the mantle source of Hawaiian volcanism of ~10⁹ yr²⁴. Given the magnitude of the elemental cerium anomalies in Molokai volcanics (>50% in sample 74-WMOL-7; F. A. Frey, personal communication), this time span would be adequate to generate cerium isotopic anomalies very well outside analytical error (>6σ) if the anomaly were the product of ancient recycling of sediment into

Table 1 Cerium isotope data

Sample	Locality	¹³⁸ Ce/ ¹³⁶ Ce	2σ	N	ε _{Ce}
74-WMOL-7	W. Molokai	1.32828	±6	1,200	-2.2
KL-2	Kiluaea, E. Rift, 1983	1.32829	±8	2,000	-2.1
2882 (NMNH 99658)	St Helena	1.32845	±9	1,200	
2926 (NMNH 99653)	St Helena	1.32844	±9	1,800	
	St Helena average	1.32844	±6	3,000	-1.0
F-33	Faial, Azores, 1958	1.32845	±8	2,700	-0.9
UPO-7	Upolu Puapua, Samoa	1.32857	±7	2,100	0.0
KG14-3	Courbet, Kerguelen	1.32856	±3	1,400	-0.1
TR-1 (NMNH 110014)	Tristan da Cunha, 1961?	1.32862	±7	4,500	+0.4

σ, Standard error; N, number of scans of mass spectrum; ε_{Ce} = parts per ten thousand deviation from bulk-earth composition calculated here.

the mantle. To test this possibility, the anomalous West Molokai basalt sample 74-WMOL-7 was analysed for ¹³⁸Ce/¹³⁶Ce, along with a 'normal' Hawaiian basalt and samples from St Helena, the Azores, Samoa, Kerguelen and Tristan da Cunha, which span the range of ¹⁴³Nd/¹⁴⁴Nd isotope ratios in OIB.

The analytical techniques adopted here for cerium isotope analysis were similar to those used by Tanaka and Masuda⁷. After separation from major elements and barium on a cation column eluted with HNO₃, very-high-purity cerium separations from the bulk rare-earth elements were made on a 30-cm-long cation column eluted with 0.25 M hydroxy-isobutyric acid (HIBA) at a pH of 4.6. Samples were cleaned to remove HIBA on a miniature cation column. Total blanks averaged 0.4 ng Ce, using reagent quantities appropriate for 0.4 g rock.

Cerium samples were loaded on tantalum-side, rhenium-centre triple filaments, and run as the oxide species on a Vg 54E mass spectrometer at East Kilbride using a single Faraday collector. Following Tanaka and Masuda⁷, the ¹⁴⁰Ce beam was usually ignored, being 'off-scale' of the amplifier, and analyses were normalized for fractionation to a ¹⁴²Ce/¹³⁶Ce ratio of 58.14. A matrix solution for interferences¹² was not considered necessary because BaO- and BaF-induced drifts in ¹³⁸Ce abundance always stopped before completion of the first 100 sets of peak scans, which were never included in the final analysis. La, Pr, Nd and Sm interferences were always monitored at first, but usually decreased to negligible levels after the first 100 scans. Interference by the peak tail of the very large ¹⁴⁰Ce¹⁶O beam was accurately corrected using a background shape interpolation, based on measurements in several different positions.

Standards run entirely at low beam current gave an average value of 464.60 ± 0.06 for the non-radiogenic ¹⁴⁰Ce/¹³⁶Ce ratio over a period of several months. This compares well with the value of 464.65 which may be calculated from Tanaka and Masuda⁷. Abundances of ¹³⁸Ce are normalized here to ¹³⁶Ce, as this yields results that are much easier to appraise than ¹³⁸Ce/¹⁴²Ce ratios.

To be sure of objectively excluding occasional data at the end of long runs when the oxide mass-fractionation law appeared to have broken down, and to check for any long-term undetected interferences, the complete run was analysed in terms of 'super-blocks' of 100 scans of the peaks. Commonly, over 20 such 'super-blocks' were available, usually from two or three separate runs, each of ~8–10 h data acquisition time. Apart from occasional deteriorations in the ratios from dying ion beams at the ends of runs, no long-term drifts in the data were observed. Quoted errors on the analyses are two standard errors on the means of the 'super-blocks'.

* Present address: Department of Geology, McMaster University, Hamilton, Ontario L8S 4M1, Canada.

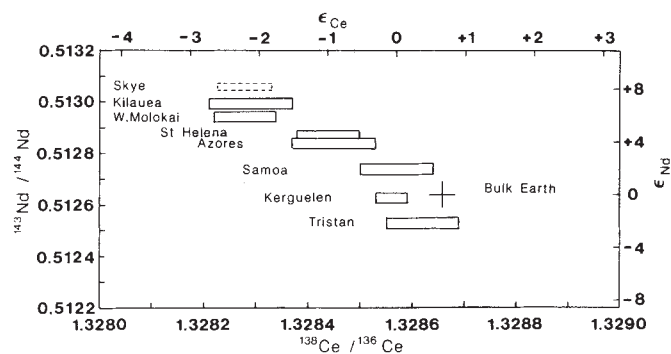


Fig. 1 Plot of $^{143}\text{Nd}/^{144}\text{Nd}$ against $^{138}\text{Ce}/^{136}\text{Ce}$ for ocean island basalts using new cerium isotope data and published neodymium isotope data^{1,13,14}, along with unpublished neodymium data for the W. Molokai sample (D. A. Clague and F. A. Frey). Bulk-Earth point from ref. 15 for comparison.

Cerium isotope ratios (Table 1) are plotted against published $^{143}\text{Nd}/^{144}\text{Nd}$ data^{1,13,14} in Fig. 1. These results define a 'mantle array' of cerium and neodymium isotopic compositions which indicates largely coherent fractionation between cerium and neodymium during mantle evolution processes. The West Molokai sample lies slightly to the left of the Ce-Nd isotope mantle array in Fig. 1, the opposite side to that predicted for incorporation of ancient subducted sediments with a negative cerium anomaly. The elemental anomaly must therefore be comparatively young. The fresh nature of the West Molokai samples rules out sub-solidus alteration, which in any case has not been observed to cause cerium anomalies in any other basalts. However, two other models cannot be easily distinguished. Clague and Frey¹¹ attributed the elemental cerium anomalies in West Molokai basalts to an unusual oxidation state in the mantle source, but they might also be attributed to the assimilation of young cerium-anomalous pelagic sediment at the base of the volcanic edifice.

A bulk-Earth $^{138}\text{Ce}/^{136}\text{Ce}$ ratio of 1.32857 is calculated from the Ce-Nd mantle array in Fig. 1. This falls within error of the bulk-Earth point calculated by Shimizu *et al.*¹⁵ from chondritic meteorites, indicating surprisingly good agreement between the East Kilbride and Japanese cerium mass spectrometry, considering the scope for systematic errors in cerium isotope measurement (for example, amplifier non-linearity for very large isotope ratios). But in view of the possibility of systematic offsets in the data, ϵ_{Ce} values represent the best method for comparison of cerium isotope data from different laboratories, and these are listed in Table 1.

Uncontaminated flood basalts from the British Tertiary Volcanic Province have strontium and neodymium isotopic compositions on the 'mantle array'¹⁶, near the field for Reykjanes Ridge basalts¹³. Therefore, in order to see whether they also fit the Ce-Nd mantle array observed here, an ϵ_{Ce} value has been calculated from isotopic data on a lava from Skye¹⁷, using a bulk-Earth La/Ce ratio of 0.38 (ref. 18) and lanthanum β -decay constant of $2.3 \times 10^{-12} \text{ yr}^{-1}$ (refs 19, 20). This sample falls on the mantle array in Fig. 1, which also attests to the overall coherency between cerium and neodymium fractionation during mantle differentiation.

Patchett *et al.*⁶ calculated that up to 1% recycling of pelagic sediment into the OIB source (2% total sediment) would be

required if this were the sole process for enrichment of this mantle reservoir. If the pelagic component consisted of material similar to "Pacific authigenic weighted mean sediment"⁹, it should (based on calculations in ref. 9, Fig. 9) generate ~50% negative cerium anomalies in the OIB source, which, after storage for 1.5 Gyr would generate cerium isotope anomalies several times the size of the analytical errors achieved here. Hence both cerium elemental and isotopic compositions preclude this model.

Nevertheless, the relatively widespread occurrence of cerium-anomalous basalts in island arcs, including the Lesser Antilles, New Guinea-New Britain, Tonga and Mariana-Izu arcs^{9,21-23}, suggests that quite substantial quantities of cerium-anomalous material may enter subduction zones. If such material subsequently enters the OIB source, but does not generate net cerium anomalies, it could be balanced by subduction of cerium-enriched manganese nodules. This would imply a uniform mixture of nodule and sediment subduction into the mantle.

Alternatively, cerium-anomalous sediment may enter subduction zones in significant quantities, but be so effectively extracted by partial melting processes below island arcs that very little reaches the deep mantle. This would imply that sediment recycling was not the principal mechanism of trace element enrichment of the OIB source.

Finally, cerium anomalies in island arc basalts might result only from the effect of subducted water on the redox conditions in the mantle melting zone. In this case the anomalies have no causal relationship with the composition of material subducted into the mantle, and no deductions about sediment recycling can be made.

More work must be done to distinguish between these alternative models in detail, but the cerium and neodymium isotope compositions of ocean island basalts demonstrate that on a broad scale these elements behave coherently during mantle differentiation.

I thank D. A. Clague, A. W. Hofmann, W. G. Melson, J. Natland and N. Watkins for providing samples, and I particularly thank F. A. Frey and W. M. White for the encouragement to proceed with this project and for selecting suitable samples. Comments by two referees were also very helpful. This work was supported by NERC until August 1985, and by the Scottish Universities.

Received 20 June; accepted 19 December 1986.

- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421-436 (1982).
- White, W. M. & Hofmann, A. W. *Nature* **296**, 821-825 (1982).
- Brown, L., Klein, J., Middleton, R., Sacks, I. & Tera, F. *Nature* **299**, 718-720 (1982).
- White, W. M., Dupre, B. & Vidal, P. *Geochim. cosmochim. Acta* **49**, 1875-1886 (1985).
- White, W. M. *Geology* **13**, 115-118 (1985).
- Patchett, P. J., White, W. M., Feldmann, H., Kielinczuk, S. & Hofmann, A. W. *Earth planet. Sci. Lett.* **69**, 365-378 (1984).
- Tanaka, T. & Masuda, A. *Nature* **296**, 214-219 (1982).
- Elderfield, H., Hawkesworth, C. J., Greaves, M. J. & Calvert, S. E. *Geochim. cosmochim. Acta* **45**, 513-528 (1981).
- Hole, M. J., Saunders, A. D., Marriner, G. F. & Tarney, J. *J. geol. Soc. Lond.* **141**, 453-472 (1984).
- Roden, M. F., Frey, F. A. & Clague, D. A. *Earth planet. Sci. Lett.* **69**, 141-158 (1984).
- Clague, D. A. & Frey, F. A. *Eos* **64**, 902 (1983).

- Nakamura, N., Tatsumoto, M. & Ludwig, K. R. *J. geophys. Res.* **89**, B438-B444 (1984).
- O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13-22 (1977).
- Newsom, H. E., White, W. M., Jochum, K. P. & Hofmann, A. W. *Earth planet. Sci. Lett.* **80**, 299-313 (1986).
- Shimizu, H., Tanaka, T. & Masuda, A. *Nature* **307**, 251-252 (1984).
- Carter, S. R., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* **281**, 28-30 (1979).
- Dickin, A. P., Jones, N. W., Thirlwall, M. F. & Thompson, R. N. *Contr. Miner. Petrol.* (in the press).
- Nakamura, N. *Geochim. cosmochim. Acta* **38**, 775-775 (1974).
- Sato, J. & Hirose, T. *Radiochem. Radioanal. Lett.* **46**, 145-152 (1981).
- Dickin, A. P. *Nature* **325**, 337-338 (1987).
- Jakes, P. & Gill, J. *Earth planet. Sci. Lett.* **9**, 17-28 (1970).
- Thirlwall, M. F. & Graham, A. M. *J. geol. Soc. Lond.* **141**, 427-445 (1984).
- White, W. M. & Patchett, P. J. *Earth planet. Sci. Lett.* **67**, 167-185 (1984).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63-87 (1978).

Early agriculture and early Postclassic Maya occupation in western Honduras

David J. Rue

Department of Anthropology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Evidence of the origins, intensification and end of cultivation can be obtained from pollen profiles indicating forest clearance. Independent archaeological data^{1,2} (the first occurrences of ceramics and village settlement) have yielded a date of ~3,000 yr BP for the onset of agriculture in western Honduras. Recent research on early Postclassic Maya demography^{3,4} has identified much regional variation in settlement densities, contrary to an earlier view, based on limited data, of immediate pan-southern Maya abandonment or depopulation after ~1,050 yr BP (ref. 5). Here I present pollen profiles which provide evidence of early slash-and-burn maize agriculture at Lake Yojoa, probably part of an Archaic hunting and gathering strategy, from ~4,500 yr BP or earlier. Slash-and-burn agriculture intensified at ~3,000 yr BP at Lake Yojoa, and stabilized thereafter. At Petapilla Swamp, 5 km north-east of the major Maya centre of Copan, the pollen sequence began at ~1,000 yr BP and suggests terminal Classic-early Postclassic (AD 950–1200) agricultural activity. The present pollen evidence of occupation until ~AD 1200 and subsequent abandonment of the Copan region reinforces recent archaeological settlement data⁶.

Western Honduras (Fig. 1) is on the southeastern periphery of the southern Maya lowlands, and is transitional to highlands in terrain, climate, and vegetation. Around the Copan Valley (610 m, 1,300 mm annual rainfall) and Yojoa Lake basin (635 m, 2,000 mm annual rainfall) rugged hills rise 1,300–2,000 m. Natural vegetation consists of mesic deciduous tropical forest on valley bottoms, upland deciduous forest on lower hills, and pine-oak montane forest on higher ridges⁷, all heavily disturbed by human occupation at present. Agricultural intensification is viable on alluvial soils, but causes erosion and nutrient depletion on slopes.

The present study of two cores collected in 1984 represents the first radiocarbon-dated (uncorrected) pollen sequences from Honduras (Fig. 2*a,b*)⁸. A 150 cm core was extracted from the southern shore of Lake Yojoa (18 km north-south, 10 km east-west, 635 m) in 50 cm of water. The gyttja sediments are increas-

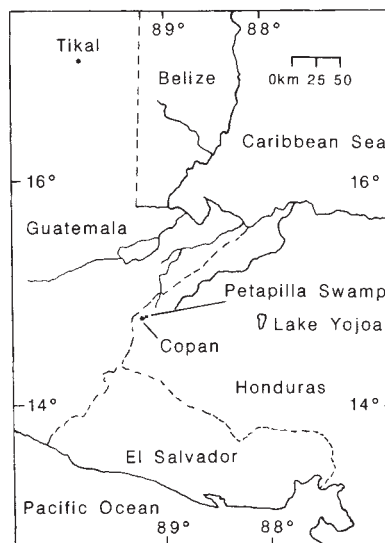


Fig. 1 Index map of sites mentioned in text.

ingly coarse below 100 cm. Burned wood fragments from 145 cm yielded a radiocarbon date of $4,770 \pm 385$ yr BP (sample UGa-5380). The presence of this wood for radiocarbon analysis minimizes hard water error encountered in dating cores elsewhere in Central America⁹. Petapilla Swamp (700 m) is a $\sim 2 \times 10^4$ m² peat bog in an intermontane basin north of the Copan Valley floor. A floating mat of vegetation (Cyperaceae, Ferns, and *Typha*) covers 2 m of water. A 135 cm core of peaty clay was extracted from shallow water at the bog's edge. Radiocarbon analysis of peat from the core yielded two dates: 595 ± 70 yr BP (sample UGa-3579) at 70 cm and 940 ± 60 yr BP (sample Beta-12789) at 130 cm.

Preservation of extant pollen is good in both cores, but there is less diversity in the palynoflora, and even lower representation of tropical deciduous forest trees than in previously reported Central American core sequences^{9–11}. Most such trees are insect-pollinated and contribute few pollen grains to sediments. There is no palynological evidence for any significant late Holocene climatic change in the present or previous Central American sequences^{9–11}, thus allowing the assumption that all vegetational changes are human-induced.

Pollen data are presented as percentages (Fig. 2*a,b*). The AP/NAP curve compares all arboreal to non-arboreal pollen.

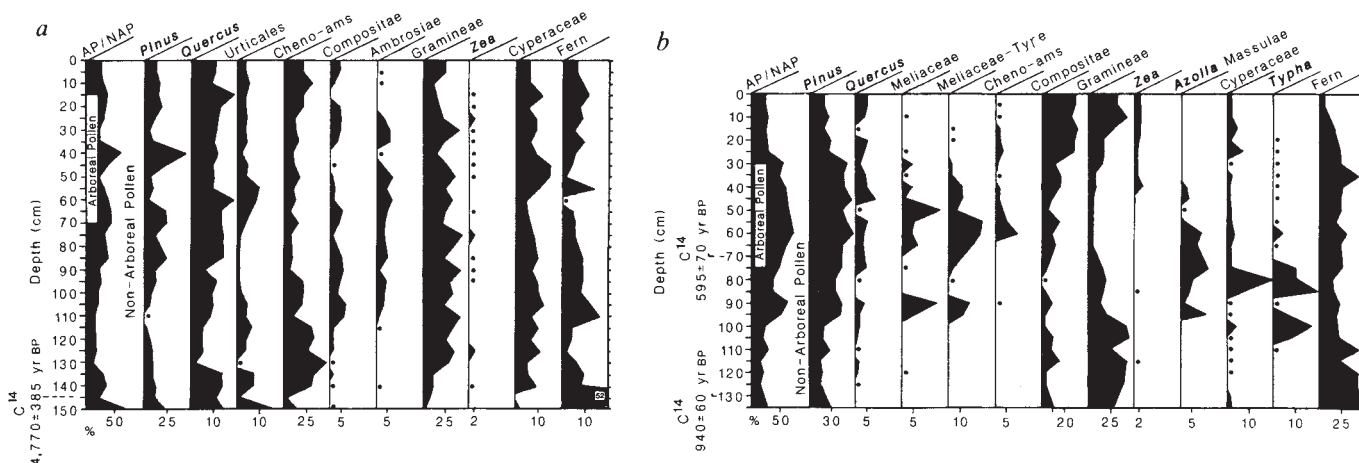


Fig. 2 *a*, A pollen diagram from Lake Yojoa, western Honduras. Analysis made at 5 cm intervals throughout. Percentages based on a pollen sum (greater than 200 grains per level) of total pollen. Dots indicate values of one per cent or less. *b*, A pollen diagram from Petapilla Swamp, western Honduras. Analysis made at 5 cm intervals throughout. Percentages based on a pollen sum (greater than 200 grains per level) of total pollen minus *Pinus* for all curves except AP/NAP and *Pinus*, which are based on total pollen. Dots indicate values of one per cent or less.

Pinus and *Quercus* are dominant members of the montane forest. Urticales, in the Yojoa profile, is an order in which identification of pollen to genus or even family is often difficult. It includes many climax (for example, *Ficus* and *Brosimum*) and successional (for example, *Trema* and *Cecropia*) tropical deciduous forest genera. Meliaceae and Meliaceae-Type, in the Petapilla sequence, are climax deciduous trees in the Mahogany family. Chenopodiaceae-Amaranthaceae ('Cheno-ams'), Ambrosiae, other Compositae, *Zea*, and other Gramineae are herbaceous taxa indicative of human disturbance. *Azolla* (an aquatic fern), Cyperaceae, and *Typha* are hydrophytes. 'Ferns' includes all Filicales.

In the Yojoa sequence (Fig. 2a), the zone from 115–150 cm relates to a period of very early slash-and-burn agriculture. Cheno-ams are dominant among herbaceous taxa, and *Zea* is present, although other Gramineae pollen percentages are low, and Ambrosiae and other Compositae are absent. Arboreal percentages decrease from the lowest levels. Disturbance is consistent with that which would be produced by long fallow slash-and-burn agriculture. Because no ceramics or village sites have been identified in Honduras before ~3,000 yr BP^{1,2}, this early clearance (~4,500 yr BP) probably occurred as part of the expanding food procurement system of preceramic late Archaic peoples, about whom little is known archaeologically^{12,13}.

Cultivation by nonceramic, principally foraging people has been identified from palaeoecological sequences elsewhere^{14,15}, and there are indications that such cultivation occurred throughout Central America as well. Phytolith analysis of sediment cores from Panama date early slash-and-burn activities to ~4,800 yr BP¹⁹, and archaeological evidence of Archaic horticulturalists has been identified tentatively in northern Belize²⁰. Previous palaeoecological studies of the Maya in which the chronology of early human disturbance was uncertain⁹ equate evidence of the earliest agriculture with the sparsely documented Swasey ceramic culture¹⁶. But as Swasey ceramics occur at a limited number of sites no earlier than 2400 BC, such disturbance was probably brought about by earlier peoples.

Agricultural intensification at Yojoa is evident above 115 cm with increases in pollen of Ambrosiae and other Compositae (field invaders). Above this point, from ~3,000 yr BP to modern times, there are few identifiable major trends. This indicates a high level of ecological stability even through the Classic period, with no evidence of vegetational changes congruent with the ~AD 900 Late Classic collapse. The maximum of *Pinus* at 40 cm, coeval with a depression in Ambrosiae and other Compositae generally, could relate to a short period of abandonment and reforestation, but such a single level occurrence is inconclusive. Populations in the Yojoa region were probably relatively static after early developments, and agriculture probably continued to be relatively extensive in nature¹⁷.

The pollen sequence from Petapilla Swamp (Fig. 2b) is short (beginning at ~1,000 yr BP), because extremely hard clays that perhaps pre-date the formation of the swamp prevented further entry of the coring device employed. The zone from 100 to 135 cm, estimated to date from ~1,000 to 750 yr BP, marks terminal Classic–early Postclassic Maya disturbance. Gramineae pollen percentages are high, *Zea* is present, and the deciduous forest trees Meliaceae and Meliaceae-Type are nearly absent. The forest was completely cleared from the valley floor and foothills during this period. Low *Pinus* percentages indicate that even montane forest may have been partially stripped. Reforestation is evident above 100 cm (after ~750 yr BP), as shown by the rise of *Pinus*, Meliaceae and Meliaceae-Type versus Gramineae. The swamp expands during this period, as shown by increased percentages of hydrophytes. Modern disturbance is present above 30 cm, with the clearance of the deciduous forest and re-emergence of Compositae and Gramineae.

The presence of *Zea* and concurring recent archaeological evidence in the form of obsidian hydration dates⁶ suggest that populations remained in the Copan region until ~750 yr BP.

This early Postclassic occupation is contrary to previous independent archaeological interpretations based solely on ceramic chronology¹⁸. No clearly differentiated ceramic style postdates Conerphase ceramics, which were assumed to end with the cessation of monumental construction ~AD 900. Early Postclassic rural populations must have continued to use utilitarian Coner styles after construction ended at the urban centre. As recent studies of other Maya regions have shown, reduced rural populations remained around generally vacant urban centres at densities which varied regionally^{3,4}. The sociopolitical and economic factors^{3,4} producing this variation are incompletely understood.

I thank the Instituto Hondureño de Antropología e Historia for research permission, Alfred Traverse for laboratory facilities, and the following for comments on a previous version of the paper: Douglas Dinsmore, William Sanders, Susan Evans, and E. S. Deevey.

Received 17 October 1986; accepted 28 January 1987.

1. Healy, P. in *The Archaeology of Central America* (eds Stone, D. & Lange, F.) 113–161 (University of New Mexico, Albuquerque, 1984).
2. Baudiez, C. & Becquelin, P. *Archéologie de los Naranjos, Honduras* (Mission Archéologique et Ethnologique Française au Mexique, 1973).
3. Chase, A. & Rice, P. *The Lowland Maya Postclassic* (University of Texas Press, Austin, 1985).
4. Willey, G. in *Late Lowland Maya Civilization* (eds Sabloff, J. & Andrews, E. V.) 17–51 (University of New Mexico Press, Albuquerque, 1986).
5. Culbert, T. P. *The Classic Maya Collapse* (University of New Mexico, Albuquerque, 1973).
6. Webster, D. & Freter, A. in *Prehistoric Maya Demography* (eds Culbert, T. & Rice, D.) (in the press).
7. Turner, B. L. II et al. in *Introducción a la Arqueología de Copan* (ed. Baudiez, C.) 35–142 (Tres Hermanos, San Jose de Costa Rica, 1984).
8. Rue, D. thesis, Pennsylvania State Univ. (1986).
9. Vaughan, H., Deevey, E. S. Jr & Garrett-Jones, S. in *Prehistoric Lowland Maya Environment and Subsistence Economy* (ed. Pohl, M.) 73–89 (Harvard University, Cambridge, 1985).
10. Cowgill, U. et al. *Mem. Conn. Acad. Arts Sci.* 17, 1 (1966).
11. Tsukada, M. & Deevey, E. S. Jr in *Quaternary Paleocology* (ed. Cushing, R. & Wright, H.) 303–331 (Yale University, New Haven, 1967).
12. Bullen, R. & Plowden, W. *Am. Antiq.* 28, 382–385 (1963).
13. Longyear, J. *Am. Antiq.* 13, 248–249 (1948).
14. Tsukada, M., Sugita, S. & Tsukada, Y. *Nature* 322, 632–634 (1986).
15. Hamilton, A., Taylor, D. & Vogel, J. *Nature* 320, 164–167 (1986).
16. Hammond, N. et al. *Am. Antiq.* 44, 92–109 (1979).
17. Boserup, E. *The Conditions of Agricultural Growth: Economics of Agrarian Change Under Population Pressure* (Aldine, Chicago, 1965).
18. Fash, W. thesis, Harvard Univ. (1983).
19. Piperno, D. *Antiquity* 59, 13–19 (1985).
20. Macneish, R. S. *Final Report of the Belize Archaeological Recon.* (Robert S. Peabody Foundation for Archaeology, Andover, 1983).

Use of non-arbitrary acoustic criteria in mate choice by female gray tree frogs

Georg M. Klump* & H. Carl Gerhardt

Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211, USA

The evolutionary significance of female choice is controversial¹. Central to the issue are the mechanisms and criteria used by females to choose among potential sexual partners. For example, females merely responding to conspicuous displays that are unrelated with male condition or quality would neither incur costs of assessment nor gain benefits over and above random mating². Here we show that selective phonotaxis by female gray tree frogs depends on non-arbitrary acoustic cues that indicate the male's species and energetic costs of signal production. Furthermore, some female preferences cannot be explained by differences in signal conspicuousness (acoustic energy) alone. Females obviously benefit by rejecting genetically incompatible, heterospecific males, which can be identified solely by pulse rate. Females also select among conspecific males on the basis of call duration, which is correlated with a male's energetic investment in courtship and probably indicates his physical condition and fitness.

* Present address: Institut für Zoologie, Technische Universität München, Lichtenbergstrasse 4, D-8046 Garching, FRG.

Table 1 Responses of female gray tree frogs in two-stimulus playback experiments with varying call duty-cycle

Test no.	Quality of alternative stimuli	Duty-cycle ratio	N	Responses	P
1	12-pulse, 900 calls per h vs 12-pulse, 1,800 calls per h	1:2	15	0:11	<0.01
2	12-pulse, 900 calls per h vs 24-pulse, 900 calls per h	1:2	15	0:12	<0.01
3	20-pulse, 1,132 calls per h vs 30-pulse, 1,132 calls per h	1:1.5	24	3:13	<0.05
4	20-pulse, 1,132 calls per h vs 40-pulse, 1,132 calls per h	1:2	11	0:9	<0.01
5	20-pulse, 607 calls per h vs 80-pulse, 607 calls per h	1:4	25	0:8	<0.01

N, number of females (one trial per female per test); pulse rate, 20 pulses per second (p s^{-1}) unless otherwise indicated; pulse duration, 25 ms, except for stimuli of 40 p s^{-1} , 12.5 ms. Probability in two-tailed binomial test shown as *P*. Each female was tested at 20°C (± 1.5) in a semi-anechoic, temperature-regulated chamber. Speakers (Analog-Digital-Systems 200) were 2 m apart; sound pressure level of alternatives was 85 dB at the release point of the female, midway between the speakers. A response was scored when a female moved to within 10 cm of a speaker, and are shown as no. of responses to the first alternative: no. of responses to the second. For details, see ref. 9.

Table 2 Responses of female gray tree frogs in two-stimulus playback experiments with minimized call duty-cycle differences

Test no.	Quality of alternative stimuli	Duty-cycle ratio	N	Responses	P
1	18-pulse, 1,000 calls per h vs 12-pulse, 1,385 calls per h	1.08:1.00	11	10:0	<0.01
2	18-pulse, 900 calls per h vs 12-pulse, 1,333 calls per h	1.01:1.00	26	9:12	NS
3	18-pulse, 900 calls per h vs 12-pulse, 1,500 calls per h	0.90:1.00	19	8:4	NS
4	24-pulse, 679 calls per h vs 12-pulse, 1,333 calls per h	1.02:1.00	39	29:7	<0.001
5	24-pulse, 638 calls per h vs 12-pulse, 1,268 calls per h	1.01:1.00	9	8:0	<0.01
6	24-pulse, 644 calls per h vs 12-pulse, 1,331 calls per h	0.97:1.00	14	5:5	NS

Symbols and conditions as in Table 1, except NS, not significant.

Table 3 Responses of female gray tree frogs in two-stimulus playback experiments with pulse rate varied

Test no.	Quality of alternative stimuli	Duty-cycle ratio	SPL ratio	N	Responses	P
1	20 pulses per s, 20-pulse, 840 calls per h vs 40 pulses per s, 40-pulse, 1,560 calls per h	1.00:1.90	1:1	13	12:1	<0.01
2	20 pulses per s, 20-pulse, 840 calls per h at 72 dB vs 40 pulses per s, 40-pulse, 840 calls per h at 90 dB	1.00:1.00	1:8	14	7:0	<0.02

Symbols and conditions as in Table 1, except that pulse rate and sound pressure levels are as shown.

Selective phonotaxis was studied by offering gravid females of *Hyla versicolor* simultaneous choices of 'synthetic sounds'. In the first set of tests, alternatives differed in either call duration or call repetition rate (call rate), but not both. Each pair of stimuli differed by at least a factor of 1.5 in terms of the call duty cycle (number of sound pulses per second of calling) and hence the amount of acoustic energy per unit time. Females invariably preferred the alternative with more sound energy (higher duty-cycle ratio), and even chose signals with combinations of duration and call rate that were more extreme than those observed in nature (Table 1, Fig. 1a, trace 3 in Fig. 1b). Halliday³ proposes that responsiveness to such 'supernormal' stimuli may be a mechanism for female choice. However, a 'supernormal' stimulus for the gray tree frog is, by definition, a 'conspicuous' signal with much more than average acoustic energy and hence an example of what Parker² terms passive attraction. Another example is the female gray tree frog's attraction to the louder of two calls^{4,5}.

In the second set of tests we varied both the duration and call rate of the alternatives so that the duty cycles of the stimuli, and hence acoustic energy, differed by 1–10% (Table 2; Fig. 1b traces 1 and 2). When duty-cycle differences favoured long calls by 2% or more, females preferred long calls to short calls (tests 1 and 4 in Table 2). When duty cycle favoured short calls, even by as much as 10%, there was no preference for either stimulus (tests 3 and 6 in Table 2). When duty-cycle differences were ~1%, females chose a 24-pulse call over a 12-pulse call but did not prefer an 18-pulse call over a 12-pulse alternative (tests 2 and 5 in Table 2). Thus, females were not simply attracted to

the stimulus pattern with more acoustic energy, but were biased toward long calls.

Taigen and Wells⁶ show that aerobic metabolism in gray tree frogs is positively correlated with call duration and call rate: their product accounts for about 84% of the variance in aerobic metabolism during calling. Wells and Taigen⁷ speculate that production of long calls depletes glycogen reserves more rapidly than production of short calls. Three additional lines of evidence indicate that long calls are more costly to produce than short calls, even when call duty-cycle differences are minimal as in the experiments just described. First, male gray tree frogs usually increase call duration in competitive situations (dense choruses or in response to playbacks of conspecific calls)⁷. Males typically compensate by lowering call rate. Note, for example, that none of the males recorded in Missouri produced calls with both the duration >850 ms and the call rate >1,000 calls h^{-1} (Fig. 1a). Second, call duration is negatively correlated with the time spent calling during an evening⁷. Third, male gray tree frogs give an exceptionally long call when they are touched by a female but do not achieve amplexus with her⁴. Males revert to baseline call duration immediately, suggesting that they cannot sustain the production of extremely long calls (G.M.K. and H.C.G., unpublished data).

So far we have discussed call properties representative of natural variation among conspecific males. In a third set of tests, we varied pulse rate, a species-specific property. Females strongly preferred synthetic calls with a pulse rate typical of conspecifics to alternatives with a higher pulse rate, in the range of a sibling species, *H. chrysoscelis*, which differs in chromosome

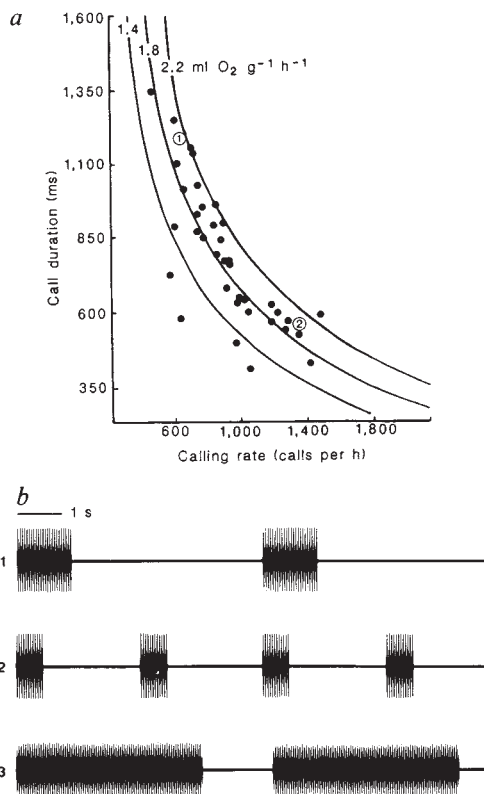


Fig. 1 *a*, Values of call duration and call rate (solid circles, calling effort) of 37 males of *Hyla versicolor* from Boone County, Missouri. Notice that if call duration was long, call rate was low. Thus, aerobic metabolism during calling (shown as solid contour lines, estimated from Wells and Taigen⁷) was usually less than 2.2 ml $O_2 g^{-1} h^{-1}$. In both Missouri and Connecticut⁷, males usually increased duration and lowered call rate when challenged by a playback of conspecific calls. *b*, Oscillograms of synthetic sounds used in playback tests of females of *H. versicolor* from Boone County, Missouri. Traces 1 and 2, 24- and 12-pulse calls respectively; trace 3, supernormal stimulus. In traces 1 and 2, the call rate of the short call was about double that of the long call; only a detailed analysis over several minutes reveals that the call duty cycle (number of pulses of 25 ms per second of calling) of the long call was ~1% greater than that of the short call. The calling effort and estimated oxygen consumption during calling that correspond to these two stimuli are indicated in *a* by the open circles. During playback tests to females (Nagra IV-S stereophonic recorder, one stimulus per channel), the timing relationship of the two alternatives was not fixed but rather the sounds drifted in and out of phase during the course of a trial. Control tests showed that the relative timing of two identical signals did not affect female preferences and that the preferences between the two signals illustrated could not be explained by the masking of one stimulus by the other. The third oscillogram illustrates a 'supernormal' stimulus; males never produced such long calls at such a high rate. The alternative to the 'super' long call was a normal 20-pulse call, the timing relationship of which was fixed so that it was always presented during the silent interval between two super long calls.

number. Preferences persisted even when the call rate of the 'conspecific' stimulus was half that of the alternative and, even more remarkably, when its sound pressure level was eight times less (18 dB) (Table 3). Thus, a female would reject a male producing an inappropriate pulse rate, regardless of the acoustic energy of his call relative to that of a conspecific male. This is a clear example of what Parker² terms active choice.

What is the evolutionary significance of our results? The (acoustic) energy-independent preference (Table 3) for sounds that differed in pulse rate, a heritable call property⁸, almost certainly reflects the fact that females mising with *H. chrysocelis* produce sterile hybrids⁹. The existence of such species recognition mechanisms is the starting point of Fisher's¹⁰ theory of sexual selection.

The preferences of the tree frogs for metabolically costly signals, as opposed to apparently arbitrary ones^{11,12}, is suggestive, but more problematic. Selective phonotaxis for louder, longer, or more rapidly-repeated calls (Table 1 and ref. 4) cannot be distinguished from passive attraction, which Parker² thinks is best viewed as a form of male-male competition because females may gain little, if anything, over and above random mating. Andersson¹³ argues that female choice for 'good genes' can evolve and be maintained. If exaggerated displays reflect a heritable component of fitness, then females exercising a choice would gain even more than the production of male offspring of greater than average attractiveness to females. We argue that energetically costly calls indicate at least that a mate is in good physical condition. Aerobic metabolism is often higher during calling than during forced exercise⁶. Moreover, the fact that females usually preferred long calls (Table 2), even when stimulus (acoustic) energy was about the same as that of short calls, indicates active choice. We predict that females using this criterion will benefit more than females which mate randomly.

We know of no data concerning heritability of call duration in gray tree frogs nor of fitness in any anuran, but recent evidence from birds¹⁴ indicates that differences in patterns of vocal production may be correlated not only with lifetime reproductive success but also survivorship. Studies of insects have docu-

mented the genetic basis of female preferences¹⁵, and frog vocalizations are polygenic traits¹⁶ for which variance is probably maintained in the face of selection¹¹. As Majerus¹⁵ points out, there is no example in which the complete system of sexual selection through female choice has been elucidated.

In summary, we have rigorously demonstrated female preferences for qualities of male displays that are not arbitrary, but rather serve to prevent hybridization or indicate directly the energetic investment of a male in courtship. These attributes of the male are encoded by different, independent parameters of the same signal, pulse rate and call duration, respectively. We argue further that the quantification of the strength of a preference, as measured by its intensity-independence, indicates the relative benefit that a female derives from choosing among males.

We thank Glenn Hausfater, Paul Harvey, Felix Breden, A. S. Rand, K. D. Wells and T. R. Halliday for valuable criticism and comments on the manuscript and Elizabeth Materna for preparing the figure. This work was supported by the Heinrich Hertz Foundation (G.M.K.), the National Science Foundation (H.C.G.) and the Research Council of the University of Missouri (H.C.G.).

Received 20 October 1986; accepted 6 January 1987.

1. Partridge, L. & Halliday, T. in *Behavioural Ecology: An Evolutionary Approach* 2nd edn (eds Krebs, J. R. & Davies, N. B.) 222-250 (Blackwell, Oxford, 1984).
2. Parker, G. A. in *Mate Choice* (ed. Bateson, P.) 141-164 (Cambridge University Press, Cambridge, 1983).
3. Halliday, T. R. in *Behavioural Ecology: An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 180-213 (Sinauer, Sunderland, Massachusetts, 1978).
4. Fellers, G. M. *Anim. Behav.* **27**, 107-119 (1979).
5. Arak, A. *Nature* **306**, 261-262 (1983).
6. Taigen, T. L. & Wells, K. D. *J. comp. Physiol. B* **155**, 163-170 (1984).
7. Wells, K. D. & Taigen, T. L. *Behav. Ecol. Sociobiol.* **19**, 9-18 (1986).
8. Burger, J. C. thesis, Univ. Missouri (1980).
9. Gerhardt, H. C. *Science* **199**, 992-994 (1978).
10. Fisher, R. A. *The Genetical Theory of Natural Selection* (Dover, New York, 1958).
11. Andersson, M. *Nature* **299**, 818-820 (1982).
12. Majerus, M. E. N., O'Donald, P. & Weir, J. *Nature* **300**, 521-523 (1982).
13. Andersson, M. *Biol. J. Linn. Soc.* **17**, 375-393 (1982).
14. Lambrechts, M. & Dohndt, A. A. *Behav. Ecol. Sociobiol.* **19**, 57-63 (1986).
15. Majerus, M. E. N. *Trends Ecol. Evol.* **1**, 1-7 (1986).
16. Gerhardt, H. C. *Behaviour* **49**, 130-151 (1974).

Genetic linkage between X-chromosome markers and bipolar affective illness

Miron Baron*, Neil Risch†, Rahel Hamburger‡, Batsheva Mandel‡, Stuart Kushner‡, Michael Newman‡, Dov Drumer‡ & Robert H. Belmaker‡§

* Department of Medical Genetics, New York State Psychiatric Institute, and Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

† Departments of Epidemiology and Public Health and of Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06510, USA

‡ Jerusalem Mental Health Center—Ezrath Nashim, and Department of Psychiatry, Hebrew University—Hadassah Medical School, Jerusalem, Israel

§ Present address: Beer-Sheva Mental Health Center, Beer-Sheva, Israel

A pedigree study shows close linkage of bipolar affective illness (manic depression) to the X-chromosome markers colour blindness and glucose-6-phosphate dehydrogenase deficiency. The maximum lod score ranges from 7.52 (assuming homogeneity) to 9.17 (assuming heterogeneity); that is, the odds in favour of linkage range between 3×10^7 to 1 and 10^9 to 1. These results provide confirmation that a major psychiatric disorder can be caused by a single genetic defect. As a possible first step in characterizing the primary genetic abnormality, this finding may have important implications for the aetiology, nosology, pathophysiology and, possibly, prevention and treatment of bipolar affective disorder. It also provides a means for identifying and characterizing homogeneous populations of patients and may help in clarifying aetiological heterogeneity.

The X-linkage hypothesis of bipolar illness has drawn considerable attention in the psychiatric and genetic literature. The hallmarks of X-linked dominant inheritance, namely, excess of affected females and lack of male-to-male transmission, have been observed in some family studies; however, instances of male-to-male transmission have also been described^{1,2}.

Linkage studies with the X-chromosome markers colour blindness, glucose-6-phosphate dehydrogenase (G6PD) deficiency and the Xg blood group have yielded conflicting results, with some pedigrees showing evidence for linkage to colour blindness or G6PD deficiency³⁻⁹, and others showing evidence against close linkage¹⁰. The conflicting data and methodological uncertainties in some of the earlier studies^{4,5} have been the subject of scrutiny¹¹. We^{1,12,13} and others¹⁴ have recently addressed these issues in a reanalysis of the linkage data. The discrepancy among the various data sets was explained by genetic heterogeneity with close linkage of bipolar illness to colour blindness and G6PD deficiency in some pedigrees; close linkage to Xg blood group was ruled out¹².

Although the diverse results can be attributed to genetic heterogeneity, the controversy has engendered calls for indepen-

dent replication using a new series of pedigrees.

Our pedigrees originated from the patient population of the Jerusalem Mental Health Center. The X-chromosome markers were colour blindness and G6PD deficiency. The loci for these traits have been localized to the distal long arm of the X chromosome (q27-qter and q28 for the colour blindness locus and *G6PD*, respectively) and are known to be closely linked^{15,16}. The population selected is particularly suitable because it is characterized by: (1) its relatively high lifetime prevalence of affective disorders generally and bipolar illness in particular¹⁷; (2) the low population rates of alcoholism and drug abuse, conditions that may confound the psychiatric diagnosis; (3) large, multigenerational pedigrees, especially in the non-Ashkenazi community¹⁸; (4) stability and accessibility of family members¹⁸; and (5) high frequency of G6PD deficiency among non-Ashkenazi Jews—as high as 60% among Kurdish males¹⁹.

The screening of probands took place between 1982 and 1985. Criteria for defining probands were: age 16 or over; a confirmed hospital diagnosis of bipolar I affective illness^{20,21}; knowledge of biological parents; and consent to participation in the research. Using the family history method, data were obtained from the proband and a first-degree relative considered to be the best informant and were supplemented by medical records. We gave preference to larger families and required that at least one immediate relative be affected with bipolar or a related major affective illness. Families with male-to-male transmission were excluded because the hypothesis of X-linkage assumes X-chromosome transmission. Families that fulfilled these initial criteria were screened for colour blindness and G6PD deficiency by obtaining information on family members from the proband and available relatives. Data on colour blindness and G6PD deficiency collected by the Israeli army on new recruits were also used, when available. Families found to segregate for either marker underwent a more extensive evaluation and these studies were extended to multigeneration pedigrees.

The probands and all available informative adult relatives (over 16) were personally examined using a slightly modified version of the Schedule for Affective Disorders and Schizophrenia (SADS)¹⁸. Diagnoses were based on the Research Diagnostic Criteria (RDC)²¹ using a diagnostic hierarchy²². A total of 85% of the relatives were personally interviewed; 5% underwent telephone interviews. Clinical data on unavailable relatives were obtained from their closest relatives using the Family History—Research Diagnostic Criteria (FH-RDC)²³. The FH-RDC method also served to supplement the personal interviews. Medical records were used when available. The clinical data were gathered by experienced mental health professionals. Good inter-rater reliability was achieved both before and during data collection; the final diagnoses were made by the clinical interviewers using all available information.

All probands met criteria for a diagnosis of bipolar I affective illness. The bipolar-related category in families of bipolar I probands reportedly consists of the following major affective disorders^{22,24}: bipolar I, bipolar II, manic (without depression),

Table 1 Lod scores for linkage of affective disorders to various markers

Marker	Pedigree no.	Ethnic origin	Recombination fraction (θ)						
			0.00	0.05	0.10	0.15	0.20	0.30	0.40
CB (Colour blindness locus)	015	NA	2.69	2.37	2.04	1.69	1.32	0.51	-0.21
	017	NA	2.31	2.11	1.90	1.68	1.45	0.94	0.40
	019	A	-4.19	-1.07	-0.69	-0.46	-0.31	-0.12	-0.02
	037	NA	1.34	1.33	1.27	1.18	1.08	0.78	0.41
Total									
All pedigrees			2.15	4.74	4.52	4.09	3.54	2.11	0.58
A pedigree			-4.19	-1.07	-0.69	-0.46	-0.31	-0.12	-0.02
NA pedigrees			6.34	5.81	5.21	4.55	3.85	2.23	0.60
G6PD	009	NA	2.94	2.67	2.38	2.07	1.75	1.07	0.37

NA, Non-Ashkenazi; A, Ashkenazi.

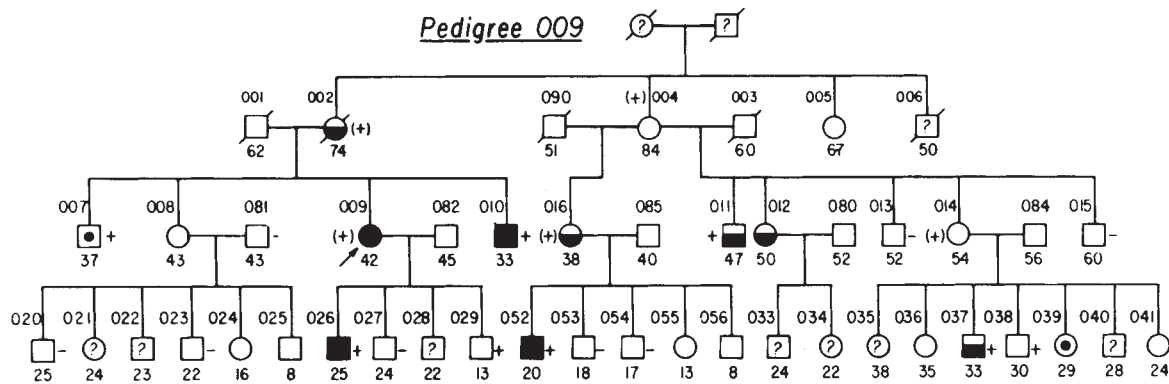


Fig. 1 Pedigree showing affective illness and G6PD deficiency. Symbols: arrow, proband; circles, females; squares, males; a slashed symbol indicates that an individual is deceased; a filled symbol denotes bipolar I, bipolar II, manic or schizoaffective (mainly affective) manic disorder; a half-filled symbol indicates unipolar or schizoaffective (mainly affective) depressive illness; and a centred dot stands for cyclothymic disorder; a blank symbol indicates other diagnoses (mainly normal) unrelated to the bipolar affective spectrum; the number above a symbol is the individual's identification number; the number below a symbol is the individual's age (in deceased individuals the age at death—if known—is indicated); +, —, and (+) denote the presence of the marker (colour blindness or G6PD deficiency), its absence, and a carrier state, respectively; ? signifies unknown psychiatric status in an adult relative (age 16 and over); individuals under age 16 were not examined.

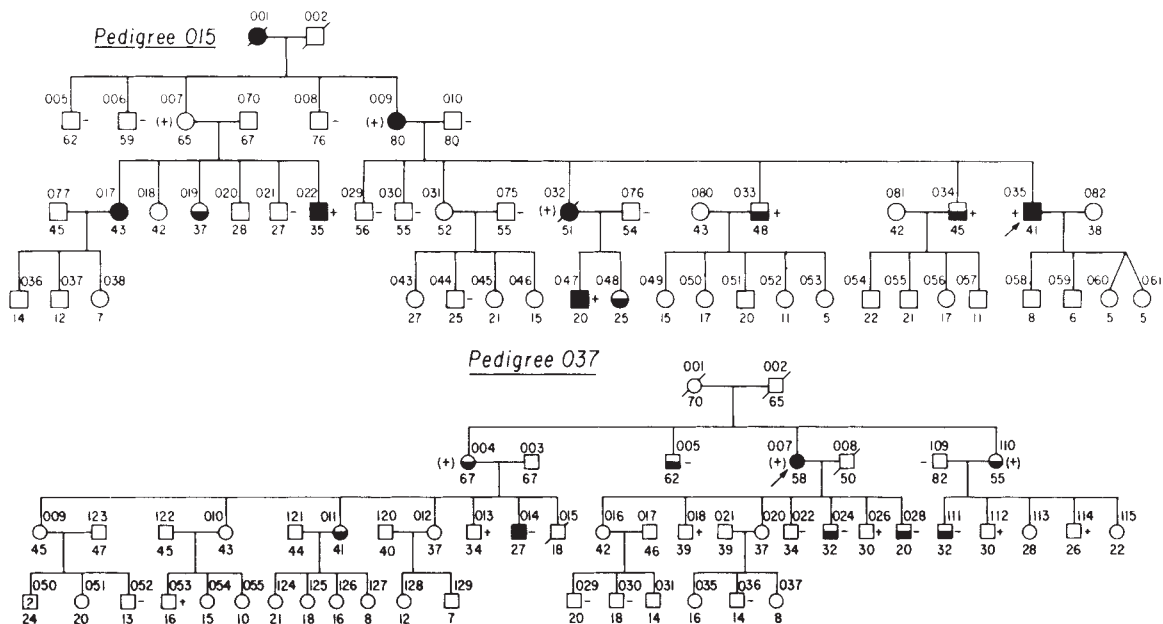


Fig. 2 Pedigrees showing affective illness and protan colour blindness. Symbols as for Fig. 1.

unipolar (major depression without mania or hypomania) and schizoaffective—mainly affective. A milder affective condition called cyclothymic disorder has also been claimed to be part of the bipolar 'spectrum'²². In accordance with the cited literature, the 'affected' category among the relatives in our pedigrees was defined by the aforementioned diagnoses.

Colour vision was examined using the Ishihara plates²⁵. Psychiatric evaluations were done independently of prior data on colour vision status; these evaluations preceded the colour vision testing in order to minimize the possibility of bias affecting the psychological assessments.

G6PD activity in erythrocytes was determined by the spectrophotometric method²⁶, the assay being done within three days of blood sampling. Split-sample and test-retest reliabilities were both high. G6PD determination and psychiatric diagnoses were always done blind to one another.

For unavailable relatives, family history for the presence of colour blindness (and the military record, if known) was accepted if there was corroboration from several relatives (one case); however, reports of absence of colour blindness (unless corroborated by the army test results) were discounted and the person's

status was considered unknown. Family history on G6PD status was accepted only if corroborated by medical records and/or the army test results (two cases); otherwise, the person's status was considered unknown.

The pedigrees that were found to be informative for the analysis of X-linkage are shown in Figs 1–3. The pedigrees contained 161 adult individuals, 47 of whom were ill with bipolar illness or related affective disorders (bipolar I, 14; bipolar II, 5; manic, 1; unipolar, 20; schizoaffective, mainly affective—manic, 2; schizoaffective, mainly affective—depressed, 2; cyclothymic, 3). There were at least 6 individuals in each pedigree with major affective illness (range: 6–11), 2 or more of whom had bipolar I, bipolar II, manic or schizoaffective—manic disorder (range: 2–8). Consanguinity was present in three pedigrees; the proband's parents in pedigrees 009 and 015 were first cousins by way of the paternal and maternal grandfathers; the proband's parents in pedigree 037 were first cousins by way of the paternal and maternal grandmothers (these portions of the pedigrees are not shown).

Four of the five pedigrees (numbers 009, 015, 017, 037) were non-Ashkenazi, whereas pedigree 019 was Ashkenazi. The non-

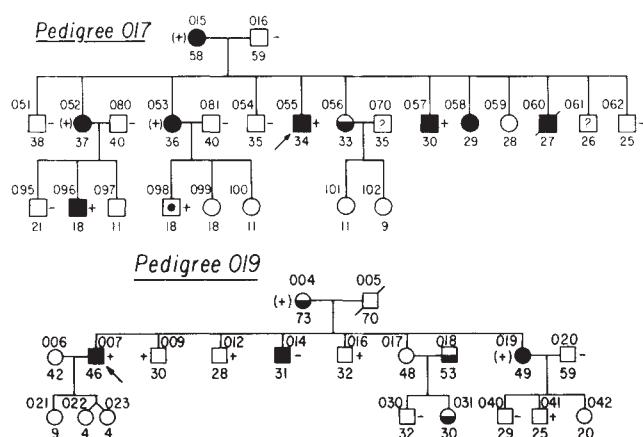


Fig. 3 Pedigrees showing affective illness and deutan colour blindness. Symbols as for Fig. 1.

Ashkenazi pedigrees originated from: Iraq (009), Yemen (015), Turkey (017) and Iran (037). The Ashkenazi pedigree originated from Poland. Pedigrees 015 and 037 segregated for protan and pedigrees 017 and 019 for deutan colour blindness, respectively. Pedigree 009 was informative for G6PD deficiency. The linkage phase in pedigrees 009, 015, 017 and 019 was coupling (that is, the alleles for the two traits—the illness and the marker—were both on the same X chromosome); the linkage phase in pedigree 037 was repulsion (the alleles for the two traits were on different X chromosomes).

The data were analysed for linkage of colour blindness or G6PD deficiency to the putative bipolar disease locus using the computer program LIPED²⁷, allowing for age-dependent penetrance²⁸ with a square-root normal age-at-onset function as described¹². In these data, mean age at onset was 19.6 years, but we used 21 years to be conservative. The data on colour blindness were combined. The frequency of the alleles for bipolar illness, colour blindness and G6PD deficiency was assumed to be 0.01, 0.04 and 0.25, respectively, based on population estimates¹⁹.

The maximum cumulative lod score (the logarithm of odds in favour of linkage) for the colour blindness pedigrees was 4.75 at recombination fraction $\theta = 0.04$; the maximum lod scores for individual pedigrees ranged from 0.00 (pedigree 019 at $\theta = 0.50$) to 2.69 (pedigree 015 at $\theta = 0.00$). The maximum lod score for the G6PD pedigree was 2.94 at $\theta = 0.00$. The lod scores are displayed in Table 1 and Fig. 4.

For the analysis of linkage heterogeneity, we used both Morton's likelihood ratio test²⁹ and an alternative, two-recombination fraction heterogeneity test¹². With either method, the degree of heterogeneity among the four colour blindness pedigrees was marginally significant: $\chi^2_3 = 7.32$, $P < 0.10$ and $\chi^2_1 = 2.91$, $P < 0.10$, respectively. All the pedigrees that support linkage had a non-Ashkenazi background whereas the only Ashkenazi pedigree had negative lod scores. The maximum lod score in the non-Ashkenazi colour blindness pedigrees is 6.34 at $\theta = 0.00$; in contrast, the corresponding lod score in the Ashkenazi pedigree at $\theta = 0.00$ is -4.19 (Table 1).

Table 2 presents a multipoint analysis for the combined lod scores for both colour blindness and G6PD. Because no pedigrees segregated both markers together, the analysis was easily done by hand. We assume the distance between the colour blindness locus and G6PD is 2 map units, or centimorgans, one per cent recombination ($\theta = 0.01$) being equal to 1 centimorgan; it is still unknown as to which marker is distal and which is proximal¹⁵. We have arbitrarily placed colour blindness at location 0 and represent G6PD as to the 'right' of it, with a positive map location (see Table 2). The inference about location of the bipolar locus and the maximum total lod score depend on the assumption of linkage heterogeneity or homogeneity among the colour blindness pedigrees. When we assume heterogeneity (excluding the single Ashkenazi pedigree), the maximum combined lod score occurs when the bipolar locus coincides with the colour blindness locus with a lod score of 9.17, or odds of 10^9 to 1 in favour of linkage. Moving in either direction from the colour blindness locus reduces the odds, for example, at 10 centimorgans on either side the odds have decreased by more than 10-fold. If we assume homogeneity, and include the Ashkenazi pedigree, the maximum lod score occurs when the bipolar locus is placed 5 centimorgans to the right of colour blindness (3 centimorgans to the right of G6PD), with a lod score of 7.52, or odds of 3×10^7 :1 in favour of linkage. The odds decrease rapidly as the possible location of the affective disorders locus moves towards the colour blindness locus, but then increase again as it moves farther away. The odds drop about 10-fold at 10 centimorgans to the right of the G6PD locus and 15 centimorgans to the left of this marker. Hence, the most probable location of the disease locus is close to and possibly between the two marker loci.

We also analysed the data using a narrower definition of the affected phenotype. When the relatives with cyclothymic disorder—individual 007 in pedigree 009 and individual 098 in pedigree 017 were classified as normal, the maximum lod scores for colour blindness and G6PD corresponding to those given above were 4.37 and 2.02. These lod scores are smaller than those obtained using the wide definition of the phenotype, but they nonetheless strongly support linkage.

In our analysis, we have assumed nearly complete penetrance for the X-linked gene at high values of age. We also analysed the data allowing for reduced penetrance and varying the gene frequency. Only small differences were obtained. Hence, the presented results are robust with regard to assumptions about gene frequency and penetrances. The results overall strongly support the existence of a genetic locus which predisposes to bipolar and related affective disorders and is closely linked to the colour blindness and G6PD loci.

This particular linkage in the population was only found in the non-Ashkenazi pedigrees, all of which showed positive lod scores. Although the number of Ashkenazi pedigrees was insufficient for comparative purposes, it is tempting to speculate that X-linked inheritance of bipolar affective illness may be more pronounced among non-Ashkenazi Jews (and, possibly, among other Mediterranean and Asian communities which share part of the same gene pool) than in other ethnic groups. Additional evidence is provided by a previous report on a large Jewish pedigree of Iranian extraction showing strong evidence

Table 2 Lod scores for location of bipolar disorders locus

Map location (centimorgans from CB) [†] [Markers]	-15	-10	-5	0 [CB]	+2 [G6PD]	+5	+10	+15
Assuming heterogeneity (non-Ashkenazi families only)	6.49	7.47	8.36	9.17*	9.07	8.55	7.71	6.74
Assuming homogeneity (all families)	6.29	6.78	7.29	4.98	7.07	7.52*	7.02	6.28

* Maximum lod score location for bipolar disorders locus.

[†] The relationship between map distance and θ is nearly linear up to $\theta = 0.15$, so from a practical standpoint, recombination frequencies can be converted directly into map distances with little loss of precision.

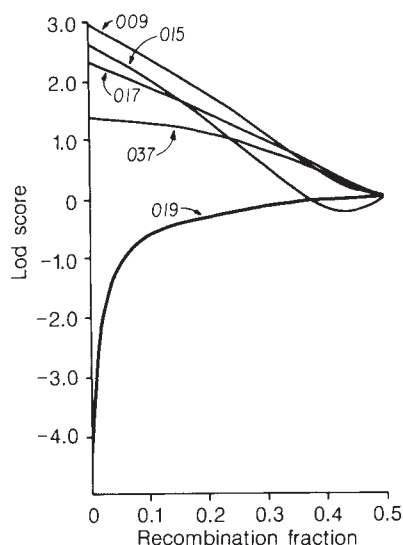


Fig. 4 Lod scores for each pedigree as a function of recombination fraction. Numbers identify the pedigree for each curve.

for linkage between bipolar illness and G6PD deficiency⁸ as well as two Sardinian pedigrees⁹.

It should be stressed that the pedigrees included in this study represent a select sample, as the study design called for the exclusion of families with male-to-male transmission. In an earlier study¹², we found significant evidence for linkage heterogeneity among previously published pedigrees, and estimated the proportion of X-linked families to be 64%. However, as pedigrees with male-to-male transmission were excluded from the analysis, that figure is not representative of the population as a whole. In a subsequent reanalysis of transmission patterns from family studies we estimated that in general perhaps one-third of the bipolar population carry the X-linked gene².

With the advent of molecular genetic techniques, many more polymorphisms have been revealed on the long arm of the X chromosome³⁰. These DNA markers, unlike the classical polymorphisms, should render most pedigrees informative for the analysis of X-linkage³⁰. This, in turn, should facilitate replication of the X-linkage findings in different populations and resolution of the question of genetic heterogeneity in the bipolar population.

We thank Drs Jurg Ott, Nancy Wexler and John Rainer for comments on the manuscript; Dr Elliot Gershon for advice and encouragement; Dr Joseph Zohar and Pnina Gilboa for helping start the project; Dr Haim Dasberg, Irwin Gordon, and Dr Bernard Lerer for administrative support; Sarah Nassi for technical help with the G6PD assays; and Sarit Baitz and Dr Zecharia Shemesh for help with clinical interviews. This work was supported by an NIMH Research Scientist Development Award MH00176 (M.B.), an NIH Research Career Development Award HD00648 (N.R.), and grants from NIMH (MH36963) and the US-Israel Binational Science Foundation (no. 3350).

Received 8 December 1986; accepted 10 February 1987.

- Baron, M., Rainer, J. D. & Risch, N. *J. Affective Dis.* **3**, 141-159 (1981).
- Risch, N., Baron, M. & Mendlewicz, J. *J. Psychiat. Res.* **20**, 275-288 (1986).
- Reich, T., Clayton, P. J. & Winokur, G. *Am. J. Psychiat.* **125**, 1358-1369 (1969).
- Mendlewicz, J., Fleiss, J. L. & Fieve, R. R. *J. Am. med. Ass.* **222**, 1624-1627 (1972).
- Mendlewicz, J. & Fleiss, J. L. *Biol. Psychiat.* **9**, 261-294 (1974).
- Baron, M. *Archs gen. Psychiat.* **24**, 721-727 (1977).
- Mendlewicz, J., Linkowski, P., Guroff, J. J. & Van Praag, H. M. *Archs gen. Psychiat.* **36**, 1442-1447 (1979).
- Mendlewicz, J., Linkowski, P. & Wilmette, J. *Br. J. Psychiat.* **137**, 337-342 (1980).
- DelZampo, M., Bocchetta, A., Goldin, L. R. & Corsini, G. U. *Acta psychiat. scand.* **70**, 282-287 (1984).
- Gershon, E. S., Targum, S. D., Matthysse, S. & Bunney, W. E. Jr *Archs gen. Psychiat.* **36**, 1423-1430 (1979).
- Gershon, E. S. & Bunney, W. E. Jr *J. psychiat. Res.* **13**, 99-117 (1976).
- Risch, N. & Baron, M. *Ann. hum. Genet.* **46**, 153-166 (1982).
- Baron, M. & Risch, N. *Neuropsychobiol.* **8**, 304-311 (1983).
- Gershon, E. S. *et al. Acta psychiat. scand.* **61**, 319-338 (1980).

- Goodfellow, P. N., Davies, K. E. & Ropers, H. H. *Cytogenet. Cell Genet.* **40**, 296-352 (1985).
- Rinaldi, A. *et al. Am. J. hum. Genet.* **30**, 339-345 (1978).
- Gershon, E. S. & Liebowitz, J. H. *Genetic Diseases Among Ashkenazi Jews* (eds Goodman, R. M. & Motulsky, A. G.) 343-352 (Raven, New York, 1979).
- Baron, M., Belmaker, R. H. & Risch, N. *Biological Psychiatry 1985* (eds Shagass, C. *et al.*) 13-15 (Elsevier, New York, 1986).
- Goodman, R. M. *Genetic Disorders Among the Jewish People* (Johns Hopkins University Press, Baltimore, 1979).
- Endicott, J. & Spitzer, R. L. *Archs. gen. Psychiat.* **35**, 837-862 (1978).
- Spitzer, R. L., Endicott, J. & Robins, E. *Archs gen. Psychiat.* **35**, 773-779 (1978).
- Weissman, M. M. *et al. Archs gen. Psychiat.* **41**, 13-21 (1984).
- Andreasen, N., Endicott, J., Spitzer, R. L. & Winokur, G. *Archs gen. Psychiat.* **34**, 1229-1235 (1977).
- Andreasen, N. C. *et al. Archs. gen. Psychiat.* (in the press).
- Ishihara, S. *Ishihara's Tests for Colour-Blindness, 24 Plates Edition* (Kanehara, Tokyo, 1981).
- Henry, R. J., Cannon, D. C. & Winkelman, D. C. (eds) *Clinical Chemistry: Principles and Techniques*, 842-847 (Harper & Row, New York, 1974).
- Ott, J. *Am. J. Hum. Genet.* **26**, 588-597 (1974).
- Hodge, S. E., Morton, L. A., Tideman, S., Kidd, K. K. & Spence, M. A. *Am. J. hum. Genet.* **31**, 761-762 (1979).
- Morton, N. E. *Am. J. hum. Genet.* **8**, 80-96 (1956).
- Drayna, D. & White, R. *Science* **230**, 753-758 (1985).

HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells

Martin Hooper*, Kate Hardy†, Alan Handyside†§, Susan Hunter† & Marilyn Monk‡

* Department of Pathology, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK

† MRC Experimental Embryology and Teratology Unit, Woodmansterne Road, Carshalton, Surrey SM5 4EF, UK

‡ MRC Mammalian Development Unit, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK

Embryonal stem (ES) cell lines, established in culture from peri-implantation mouse blastocysts¹⁻³, can colonize both the somatic and germ-cell lineages of chimaeric mice following injection into host blastocysts^{4,5}. Recently, ES cells with multiple integrations of retroviral sequences have been used to introduce these sequences into the germ-line of chimaeric mice⁵, demonstrating an alternative to the microinjection of fertilized eggs⁶ for the production of transgenic mice. However, the properties of ES cells raise a unique possibility: that of using the techniques of somatic cell genetics to select cells with genetic modifications such as recessive mutations, and of introducing these mutations into the mouse germ line. Here we report the realization of this possibility by the selection *in vitro* of variant ES cells deficient in hypoxanthine guanine phosphoribosyl transferase (HPRT; EC 2.4.2.8), their use to produce germline chimaeras resulting in female offspring heterozygous for HPRT-deficiency, and the generation of HPRT-deficient preimplantation embryos from these females. In human males, HPRT deficiency causes Lesch-Nyhan syndrome⁷, which is characterized by mental retardation and self-mutilation.

The structural gene for HPRT is X-linked both in mice⁸ and in humans⁷. Thus, male cells are hemizygous and HPRT⁻ variants, readily selected by resistance to the purine analogue 6-thioguanine⁹, arise at relatively high frequency. Although HPRT⁻ embryonal carcinoma (EC) cells colonized the somatic tissues of chimaeric mice, they failed to colonize the germ line¹⁰. Failure to colonize the germ line is a common limitation with EC cell lines, although the problem with the HPRT⁻ cells may have been compounded by the use of a chemical mutagen. Also, selection in the presence of feeder cells causes the loss of resistant variants by transfer of 6-thioGMP from wild-type parental cells through the feeder layer¹¹ and probably exposes surviving HPRT⁻ cells to sublethal levels of 6-thioGMP which are mutagenic¹². We therefore chose to work with an embryo-derived ES cell line, to isolate spontaneous HPRT⁻ variants,

§ Present address: Institute of Obstetrics and Gynaecology, Hammersmith Hospital, Du Cane Road, London W12 0HS, UK.

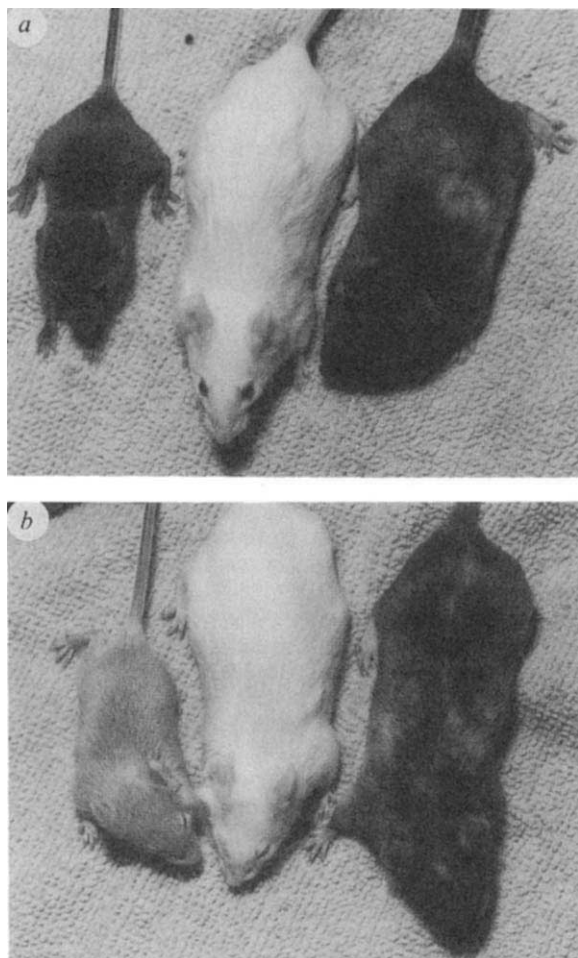


Fig. 1 Male E14TG2a chimaeras and their offspring showing the transmission of coat colour markers. *a*, Male chimaera (right), the BALB/c female to which it was mated (centre) and a host blastocyst-derived offspring (left). *b*, One of the five male chimaeras with complete colonization of the germ line (right), the BALB/c female to which it was mated (centre) and one of the E14TG2a-derived offspring (left). BALB/c mice are homozygous for the albino allele (c/c), so that offspring derived from host sperm are normally pigmented (C/c), but those derived from E14TG2a sperm are pale chinchilla (c^{ch}/c).

Methods. F1 (C57B1/6JLac $\times$ CBA/CaLac) female mice (3 to 4 week old) were superovulated by intraperitoneal injection of 7.5 IU pregnant mares serum gonadotrophin at 12:00 followed 48 h later by 5 IU human chorionic gonadotrophin (Intervet), paired with males of the same strain, and examined the following day for copulation plugs (day 1 of pregnancy). Morulae and early blastocysts were flushed from the oviducts and uteri late on day 3 in HEPES-buffered medium M2²³ supplemented with 0.4 mg ml⁻¹ bovine serum albumin (BSA) and cultured overnight to the expanded blastocyst stage in medium M16²⁴ supplemented with 0.4 mg ml⁻¹ BSA at 37 °C in an atmosphere of 5% CO₂ in air. Cells were grown on gelatin-coated tissue-culture plasticware (Lux; Flow Laboratories) in BRL-conditioned medium with 20% (v/v) serum (30 ml medium CM β) incubated for three days with a confluent 175-cm² monolayer of BRL cells²⁵, filtered through a membrane of pore size 0.2 μ m, and then diluted with 15 ml fresh medium CM β and an additional 5 ml fetal calf serum (FCS) at 37 °C in an atmosphere of 95% air, 5% CO₂. Medium CM β is Glasgow-modified Eagle's medium with 1 $\times$ non-essential amino acids and 1 mM sodium pyruvate (Flow Laboratories), 10% (v/v) FCS (Northumbria Biologicals, Cramlington, Northumberland) and 0.1 mM β -mercaptoethanol. Cells were subcultured twice weekly with 0.25% (w/v) trypsin, 1 mM Na₂EDTA, 1% (v/v) chicken serum and monitored periodically for mycoplasma contamination by the method of Chen²⁶. No such contamination was detected during this study. Colonies of cells were selected with a micropipette, washed twice in Hanks Balanced Salts solution without calcium and magnesium ions (Imperial Labs), incubated in 0.125% (w/v) trypsin, and finally washed and disaggregated in medium CM β with 20% (v/v) FCS. Cells were injected into blastocysts using a De Fonbrune micromanipulator and Agla microsyringe. Injected blastocysts were allowed to re-expand in culture for 1 to 2 h before transfer to the uteri of day-3 pseudopregnant recipients.

and to avoid feeder layers during the selection procedure by the use of Buffalo rat liver cell (BRL)-conditioned medium which inhibits differentiation *in vitro*^{11,13}.

The ES cell line, E14, was derived from strain 129/Ola mouse blastocysts by trypsinizing immunosurgically isolated inner cell masses and culturing the dissociated cells on STO fibroblast feeder cells in BRL-conditioned medium (A.H.H., K.H., M. Jones and M.L.H., manuscript in preparation). Following expansion into mass culture, E14 cells were maintained in BRL-conditioned medium alone. For the isolation of HPRT⁻ variants, 4.4×10^6 E14 cells were seeded at 7×10^3 cells cm⁻² in BRL-conditioned medium with 20% serum (see Fig. 1 legend) on gelatin-coated dishes. The cells attached overnight and then the medium was replaced with selective medium containing 10 μ g ml⁻¹ 6-thioguanine which was changed every 2–3 days. After 13 days, seven surviving colonies were visible. By picking and subculturing in non-selective medium, sublines of undifferentiated cells were established from four of these colonies. These were screened for thioguanine resistance, sensitivity to HAT (hypoxanthine, aminopterin and thymidine)¹⁴, karyotype and ability to differentiate *in vitro*, and the subline E14TG2a was chosen for further study. Like E14 cells, E14TG2a cells had a modal chromosome number of 40 (16 of 20 spreads had the modal number). They were identified as male by the presence of a small uniformly fluorescent chromosome, identified as the Y chromosome¹⁵, after staining with bis-N, N' (6-chloro-2-methoxy-acridin-9-yl)spermidine ((CMA)₂S) (which gives similar staining to the spermine derivative¹⁶). Further, E14TG2a cells differentiated extensively *in vitro* following formation of suspension aggregates in non-conditioned medium

and subsequent outgrowth (ref. 13; and A.H.H. *et al.*, in preparation). However, E14TG2a cells were resistant to 10 μ g ml⁻¹ 6-thioguanine at cell densities at which E14 cells were killed by less than 0.1 μ g ml⁻¹ of this analogue, and unlike E14 cells they failed to grow in HAT medium. This suggested a deficiency in HPRT, and indeed when this enzyme was assayed directly by the method of Gillin *et al.*¹⁷ no activity was detectable (specific activities: E14, 0.57 ± 0.06 (mean $\pm$ s.e.m.) pmol inosinic acid (IMP) per min per μ g cell protein; E14TG2a < 0.002 pmol IMP per min per μ g cell protein).

Between 5 and 20 HPRT⁻ E14TG2a cells were injected into each of a series of host F2 (C57B1/6JLac $\times$ CBA/CaLac) blastocysts. Of 396 injected blastocysts transferred to recipients which became pregnant, 266 mice (67%) were born. Chimaeras were identified by examining the coat colour of litters at 7–10 days of age. F2 host blastocysts are homozygous for the wild-type allele at the c locus (C/C), but 129/Ola-derived E14TG2a ES cells are homozygous for the chinchilla allele (c^{ch}/c^{ch}) which produces a lighter coat colour. If the coat was not chimaeric, blood samples were taken at about 4 weeks for glucose phosphate isomerase isozyme (GPI)⁴ analysis (F2 host blastocyst: $Gpi-1^b/Gpi-1^b$; 129/J-derived E14TG2a ES cells: $Gpi-1^a/Gpi-1^a$). However, blood isozyme analysis of 119 of these mice failed to detect any more chimaeras and so this additional screening was discontinued. The incidence of chimaerism ranged from 0 to 38% between experiments, but overall a total of 63 (24%) chimaeras were born. This is comparable with the incidence obtained with other ES lines⁴ although lower than recently achieved with one particular line⁵.

The distortion of the sex ratio and the pattern of germline

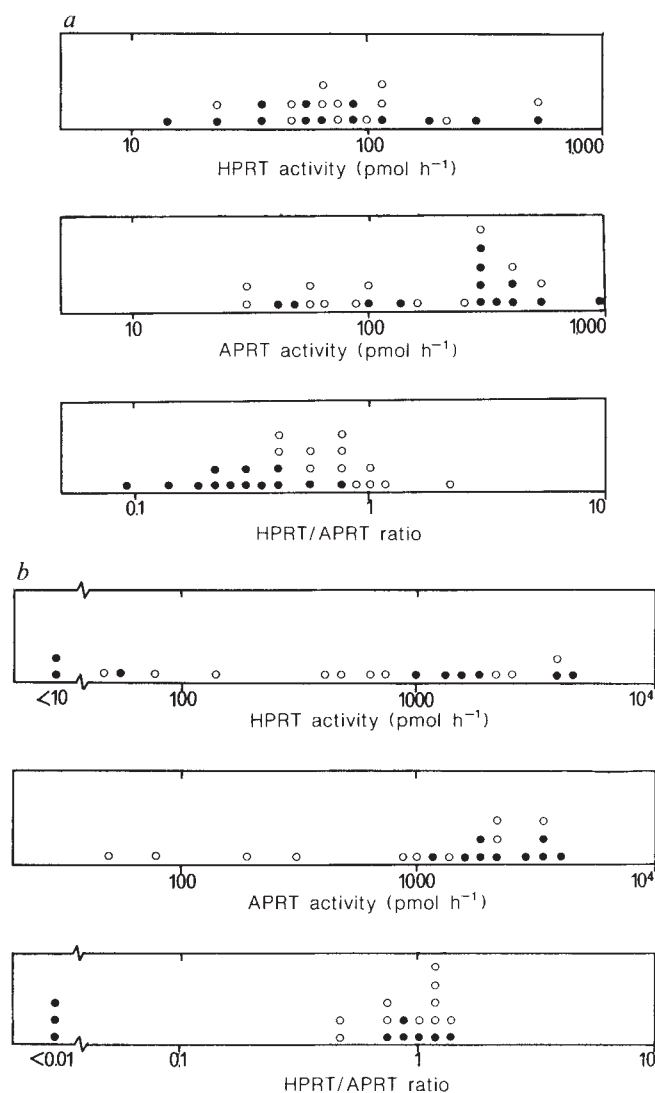


Fig. 2 Distribution of HPRT and APRT activities and their ratios in *a*, single hair follicles and *b*, single intestinal crypts. ●, Presumptive *Hprt*⁺/*hprt*⁻ heterozygous females, and ○, controls (male E14TG2a-derived littermate and BALB/c female in (*a*) and BALB/c female in (*b*)). Each data point represents a single follicle or crypt. The vertical co-ordinate of individual data points has no significance other than to avoid superimposition.

Methods. Hairs with intact follicles, and intestinal crypts isolated as described¹⁹ were placed in 10-μl capillary tubes with approximately 5 μl medium M2 supplemented with 0.4 mg ml⁻¹ polyvinylpyrrolidone, the column of fluid shaken into the centre and the ends heat-sealed. Samples were freeze-thawed three times in liquid nitrogen and HPRT and APRT activities assayed as described²⁷.

transmission is similar to that observed with other male ES cell lines^{4,5}. Fifty chimaeras are phenotypically male, and 11 female, by external examination (two mice died before sexing). To date, 19 of 34 male chimaeras test bred with females carrying appropriate markers have shown germline transmission of the *c*^{ch} marker derived from the HPRT⁻ E14TG2a ES cells (Fig. 1). With six males the marker is transmitted to all the offspring, and among the other thirteen the proportion of E14TG2a-derived offspring ranges between 1 and 56% and in many cases is higher than previously observed with other lines for partial transmission⁴.

On the assumption that the HPRT-deficiency of E14TG2a cells is due to an X-linked structural gene mutation, all female offspring carrying the *c*^{ch} marker should be heterozygous for HPRT-deficiency and all male offspring wild-type. Human female carriers for Lesch-Nyhan syndrome are identified by

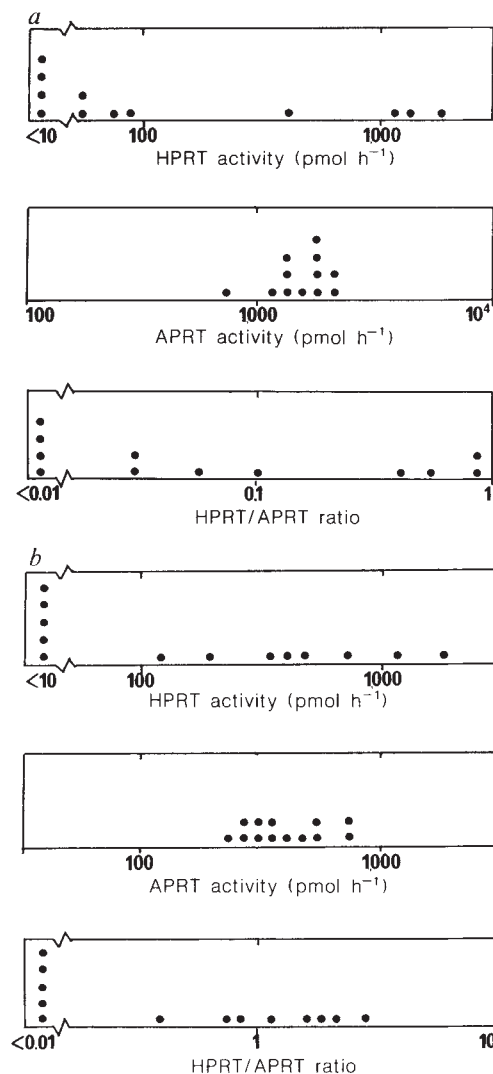


Fig. 3 Distribution of HPRT and APRT activities and their ratios in single day-3 morulae from matings between heterozygous females and wild-type F1 (C57B1/6JLac x CBA/CaLac) males: *a*, litter no. 1; *b*, litter no. 2. Each data point represents a single morula. Plotting of points is as in Fig. 2.

determining the ratio of activities of HPRT and the autosomal enzyme, adenine phosphoribosyl transferase (HPRT/APRT ratio), in single hair follicles. Follicles from a single heterozygote are heterogeneous with respect to HPRT/APRT ratio, with normal, completely HPRT-defective and intermediate phenotypes present¹⁸, presumably reflecting the mosaic distribution of HPRT⁻ cells in the scalp resulting from random X-inactivation. Although we do not see completely HPRT⁻ follicles in the heterozygous female E14TG2a-derived offspring, presumably because some HPRT⁺ cells are present in all our samples, the HPRT/APRT ratios show a distribution of values lower than those for wild-type males and females (Fig. 2*a*). Moreover, in a representative heterozygous female, some intestinal crypts, each of which is a clonal population of cells with respect to X-inactivation¹⁹, have undetectable HPRT activity, although the APRT activities fall within the control range (Fig. 2*b*). These results are consistent with the mosaic distribution of HPRT⁻ cells which would be expected in females generated from spermatozoa carrying an X-linked *hprt*⁻ mutation, and confirm the germline transmission of the mutant allele. Germline transmission of an *hprt*⁻ allele derived from an ES cell line subjected to retroviral insertion mutagenesis has been independently achieved by Kuehn *et al.*²⁰.

Heterozygous females were mated with normal males and their preimplantation embryos assayed at the morula stage. Some of the embryos from each of two such litters showed undetectable levels of HPRT activity (Fig. 3) as would be expected for hemizygous male embryos carrying the mutant allele on their single X chromosome, indicating that the mutant allele can be transmitted through the female germ line and the resulting HPRT-deficient embryos survive to the morula stage. Nine litters from similar matings have been allowed to develop to term. The litter sizes at birth ranged from 8 to 13, with 52 males and 40 females. On the basis of enzyme assay on blood samples from 41 of these mice, 11 out of 21 males and none of 20 females are HPRT-deficient. There is therefore no significant loss of HPRT-deficient male embryos *in utero*. No postnatal deaths have yet occurred in these litters, which are now between three and five weeks old.

In humans, HPRT-deficient males exhibit a characteristic pattern of behavioural abnormalities (Lesch-Nyhan syndrome), of which the principal features are mental retardation, spastic cerebral palsy, choreoathetosis, and self-mutilative biting⁷. There is circumstantial evidence that this is caused by the development of lesions in the dopamine-mediated neural pathways of the basal ganglia of the brain²¹, but the absence of any known case of HPRT-deficiency in an experimental animal has so far prevented a definitive test of this hypothesis. This may now be possible with these HPRT-deficient mice. Although no overt behavioural abnormalities have been noted either in HPRT⁻ progeny of heterozygous females (up to their present age of five weeks) or in male chimaeras with apparently extensive HPRT⁻ cell contributions, detailed behavioural investigations have yet to be carried out. HPRT-deficient mice may also be valuable for evaluating therapeutic procedures such as chronic L-DOPA therapy or neural transplants²¹, and as a model system for rescuing defective embryos by cell replacement or gene therapy²². In addition, the use of HPRT-deficient mouse embryos will facilitate the development of techniques for the preimplantation diagnosis of Lesch-Nyhan syndrome and other inherited diseases using biopsies of early human embryos. Finally, the general approach described here for introducing genetic modifications into the mouse germ line be applicable to other genetic lesions of developmental and clinical interest.

We thank Mary Jones for technical assistance and George Spowart, MRC Clinical and Population Cytogenetics Unit, Edinburgh, for a gift of (CMA)₂S. This work was partly supported by a grant from the Cancer Research Campaign to M.L.H.

A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice

Michael R. Kuehn, Allan Bradley,
Elizabeth J. Robertson & Martin J. Evans

Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK

The human Lesch-Nyhan syndrome is a rare neurological and behavioural disorder, affecting only males, which is caused by an inherited deficiency in the level of activity of the purine salvage enzyme hypoxanthine-guanosine phosphoribosyl transferase (HPRT)¹⁻³. How the resulting alterations in purine metabolism lead to the severe symptoms characteristic of Lesch-Nyhan patients is still not understood^{3,4}. No mutations at the *Hprt* locus leading to loss of activity have been described in laboratory animals. To derive an animal model for the Lesch-Nyhan syndrome, we have used cultured mouse embryonic stem cells⁵, mutagenized by retroviral insertion and selected for loss of HPRT activity, to construct chimaeric mice^{6,7}. Two clonal lines carrying different mutant *Hprt* alleles have given rise to germ cells in chimaeras, allowing the derivation of strains of mutant mice having the same biochemical defect as Lesch-Nyhan patients. Male mice carrying the mutant alleles are viable and analysis of their cells shows a total lack of HPRT activity.

To increase the frequency of mutations in the X-linked *Hprt* gene in a male embryonic stem cell line, multiple infection with a retroviral vector was carried out^{7,8}. HPRT⁻ variants were isolated from infected stem cell populations using 6-thioguanine (6-TG)⁹. From a total of 10⁷ originally infected cells, four 6-TG-resistant clones were isolated, TG^R4, 5, 6 and 7. In a separate experiment an additional 6-TG-resistant clone, TG^R3, was recovered from 5 × 10⁶ originally infected cells. Selections of 10⁸ uninfected cells have not yielded any HPRT⁻ clones. All five of these clones were judged to be HPRT⁻, in that each of the 6-TG-resistant clonal lines was killed within three days in HAT (hypoxanthine, aminopterin and thymidine) medium¹⁰ but showed equivalent plating efficiencies in 6-TG medium and normal medium.

Southern analysis of DNA prepared from each of the mutant clones revealed alterations in the restriction fragment pattern of the *Hprt* gene¹¹, as expected if gene disruption by retroviral integration had occurred. Using enzymes that cut within the locus, an identical altered pattern was seen for the *Hprt* gene in clones TG^R4, 5, 6 and 7. This alteration is in the 3' region of the gene between exons 5 and 6 (Fig. 1). Reprobing with the *neo* gene to determine the integration pattern of the viral vector has shown that these four clones are derived from one mutated cell, the daughter cells of which had been infected several more times before 6-TG selection. Southern analysis of the *Hprt* gene in TG^R3 has revealed an alteration in the 5' region of the gene within the *Hprt* locus but does not cut the vector. *Kpn*I digestion of TG^R3, 4 or 6 DNA produces two novel fragments rather than within the *Hprt* locus but does cut the vector. *Kpn*I digestion of TG^R3, 4 or 6 DNA produces two novel fragments rather than the one fragment of 50 kilobases (kb) seen with wild-type DNA (Fig. 1D). The appearance of two novel fragments implies that a virus has integrated within the gene thereby introducing *Kpn*I sites into the locus (Fig. 1E). This analysis supports the conclusion that these clones are HPRT⁻ as a result of insertional mutation.

Cells from three of the HPRT⁻ clonal lines, TG^R3, 4 and 6, were used to produce chimaeric animals (see Table 1). The contribution of the injected stem cells to somatic tissues of the chimaeras was extensive and uniformly high. The expected

Received 30 December 1986; accepted 3 February 1987.

- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154-156 (1981).
- Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634-7638 (1981).
- Silver, L. M., Martin, G. R. & Strickland, S. (eds) *Teratocarcinoma Stem Cells* (Cold Spring Harbor Laboratory, New York, 1983).
- Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. *Nature* **309**, 255-256 (1984).
- Robertson, E., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445-448 (1986).
- Gordon, J. W. & Ruddle, F. H. *Gene* **33**, 121-136 (1985).
- Nyhan, M. L. *A. Rev. Med.* **24**, 41-60 (1973).
- O'Brien, S. J. (ed.) *Genetic Maps* Vol. 2 (National Institutes of Health, Bethesda, Maryland, 1982).
- Caskey, C. T. & Kruh, G. D. *Cell* **16**, 1-9 (1979).
- Dewey, M. J., Martin, D. W., Martin, G. R. & Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5564-5568 (1977).
- Hooper, M. L. in *Teratocarcinomas: a Practical Approach* (ed. Robertson, E.) (IRL, Oxford, in the press).
- Engel, W., Krone, W. & Wolf, U. *Mut. Res.* **4**, 353-368 (1967).
- Smith, A. & Hooper, M. L. *Dev. Biol.* (in the press).
- Littlefield, J. W. *Science* **145**, 709-710 (1964).
- Nesbitt, M. N. & Francke, U. *Chromosoma* **41**, 145-158 (1973).
- Van de Sande, J. H., Lin, C. C. & Deugau, K. V. *Expl. Cell Res.* **120**, 439-444 (1979).
- Gillin, F. D., Roufa, D. J., Beaudet, A. L. & Caskey, C. T. *Genetics* **72**, 239-252 (1972).
- Gartler, S. M., Scott, R. C., Goldstein, J. L. & Campbell, B. *Science* **172**, 572-573 (1971).
- Ponder, B. A. J. *et al. Nature* **313**, 689-691 (1985).
- Kuehn, M. R., Bradley, A., Robertson, E. J. & Evans, M. *Nature* **326**, 295-298 (1987).
- Baumeister, A. & Frye, G. D. *Neurosci. Behav. Rev.* **9**, 169-178 (1985).
- Patel, P. I. & Caskey, C. T. *Bioessays* **2**, 4-8 (1985).
- Quinn, P., Barros, C. & Whittingham, D. G. *J. Reprod. Fert.* **66**, 161-168 (1982).
- Whittingham, D. G. *J. Reprod. Fert. Suppl.* **14**, 7-21 (1971).
- Coon, H. G. *J. Cell Biol.* **39**, 29a (1968).
- Chen, T. R. *Expl. Cell Res.* **104**, 255-262 (1977).
- Monk, M. & Kathuria, H. *Nature* **270**, 599-601 (1977).

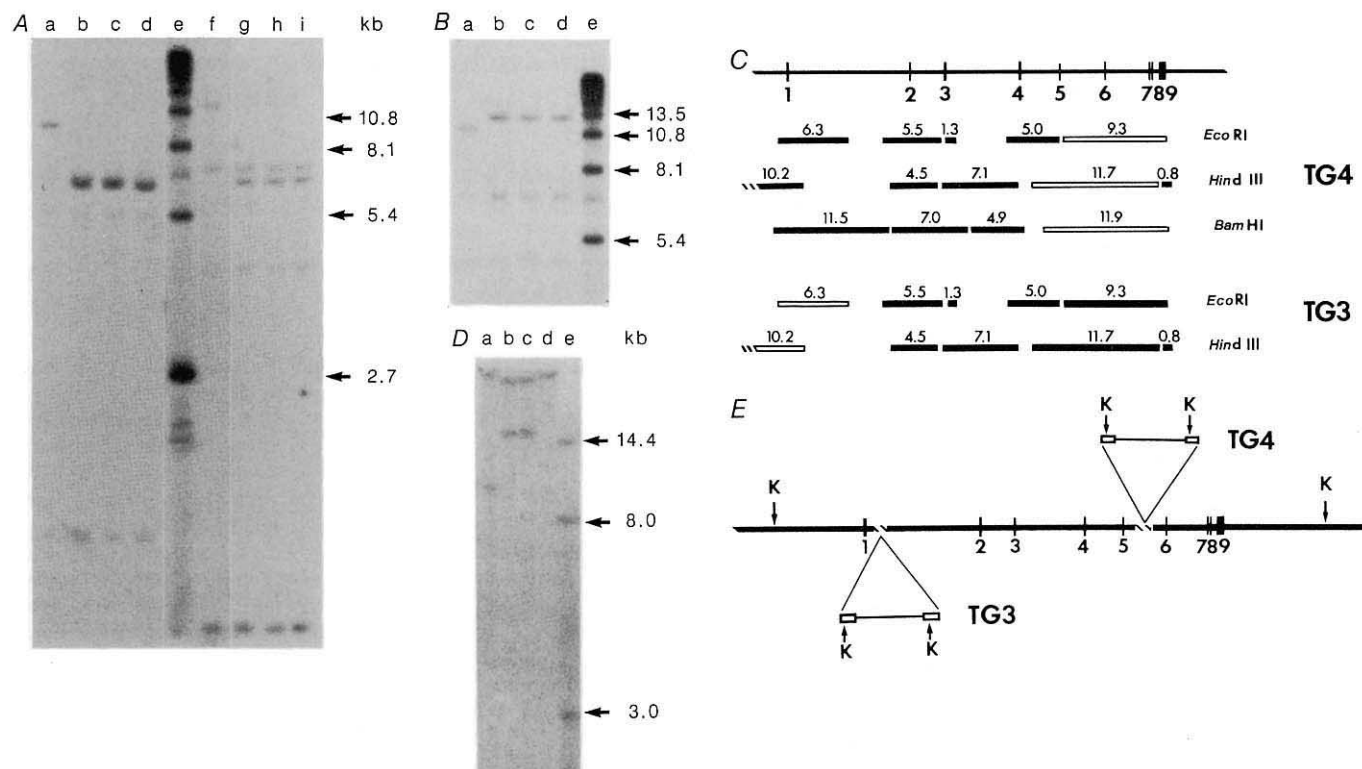


Fig. 1 Southern analysis of 6-TG-resistant clonal lines. DNA samples ($\sim 3 \mu\text{g}$), prepared from stem cells passed twice on gelled plates to reduce feeder-cell numbers, were restriction digested, fractionated on agarose gels, transferred to nylon membranes, hybridized with nick-translated pHPRT²⁷, washed and autoradiographed using standard procedures. **A**, *EcoRI* and *HindIII* analysis is shown in lanes a–d and f–i respectively. Lanes a and f, wild-type DNA; lanes b and g, clone TG^R4; lane c, clone TG^R5; lanes d and h, clone TG^R6; lane i, clone TG^R7. The sizes of the bands of the size marker lane (e) are shown at the side in kb. **B**, *BamHI* analysis. Lane a, wild-type DNA; lane b, TG^R4; lane c, TG^R5; lane d, TG^R7; lane e, size marker, sizes are shown in kb beside the lanes. **C**, map of the wild-type mouse *Hpirt* gene¹¹. Exons are numbered sequentially from the 5' end of the gene. Fragments produced by the restriction enzymes used in this Southern analysis and that hybridize to the full-length HPRT cDNA probe are shown below the map as dark boxes. The signal strength of some fragments, such as the 6.3-kb *EcoRI* fragment, is low in the blot shown above owing to the small size of the exon contained within the fragment. The open boxes refer to wild-type restriction fragments which are missing in DNA from clones TG^R4, 5, 6 and 7 (labelled TG4). The fragments missing in DNA from clone TG^R3 are shown as open boxes below the map for TG^R3, 4, 5, 6 and 7. **D**, *KpnI* analysis of DNA from the TG^R3, TG^R4 and TG^R6 clonal lines to determine retroviral integration into the *Hpirt* gene. Lane a, TG^R3 DNA; lane b, TG^R4 DNA; lane c, TG^R6 DNA; lane d, wild-type DNA; lane e, *PstI* digest of wild-type DNA to provide size markers (sizes shown at the side). **E**, retroviral integration into the *Hpirt* locus introduces novel *KpnI* sites (designated K) into the gene, producing two *KpnI* restriction fragments which hybridize to the cDNA probe in a Southern analysis. As shown in **D**, wild-type DNA has only one 50-kb *KpnI* fragment whereas TG^R3, 4 and 6 have two *KpnI* fragments, the expected result of viral integration.

distortion in the overall sex ratio was observed in the TG^R3 and TG^R6 series of chimaeras⁶, but was lacking in the population of chimaeras produced using TG^R4 cells. Karyotype analysis of the TG^R4 clonal line showed that 20 out of 41 mitoses scored by G-banding¹² lacked a Y chromosome, whereas TG^R3 and TG^R6 were normal. As XO cells in a female mouse can give rise to gametes¹³, the female chimaeras of the TG^R4 series as well as all the male chimaeras were test bred. TG^R4-derived cells have colonized the germ line of one female and seven male chimaeras. TG^R3-derived cells have contributed to the germ line of three male chimaeras (see Table 1).

Southern analysis of DNA prepared from five of the female progeny of the female germline chimaera, TG.4.52, has shown that two of these F₁ animals have a restriction fragment pattern diagnostic for both the altered allele and the normal *Hpirt* gene. The other three F₁ females have only the normal fragment pattern (Fig. 2). Karyotype analysis of primary fibroblast cultures derived from the F₁ females has shown that the animals with only the normal *Hpirt* pattern have a 39, XO chromosomal constitution. In addition to showing that the altered allele has been established in the mouse germ line, this analysis demonstrates that the germ cells of the TG.4.52 chimaera are derived from XO TG^R4 cells. Southern analysis of female F₁ animals sired by the germline chimaeras of the TG^R3 series has demon-

strated that this allele has also been established in the germ line (data not shown). This allele has been named *Hpirt-b*^{m1} and the allele carried by the TG^R4-derived animals has been named *Hpirt-b*^{m2}.

To determine whether these two newly established *Hpirt* alleles continue to cause a mutant phenotype the primary fibroblast cultures derived from the heterozygous F₁ females were further analysed. Owing to random X-chromosome inactivation a percentage of these heterozygous cells should lack HPRT activity^{14,15}. After two weeks in 6-TG medium, heterozygous cells from animals bearing either allele were not impaired in growth. Control fibroblasts derived from albino littermates and from the XO F₁ animals carrying only the normal X-chromosome were killed within one week. As a further assay of HPRT activity, unselected primary fibroblasts were plated at low density in medium containing ³H-hypoxanthine^{14,15}. Incorporation of the labelled compound into nucleic acids is dependent on HPRT activity. Autoradiography has shown that whereas 100% of the HPRT⁺ control cells were found to be labelled, approximately 40% of the cells derived from the heterozygous F₁ animals were unlabelled (Fig. 3). This and the first test confirm that cells expressing these altered *Hpirt* alleles are deficient in HPRT activity.

To derive hemizygous mutant males, the F₁ females heterozy-

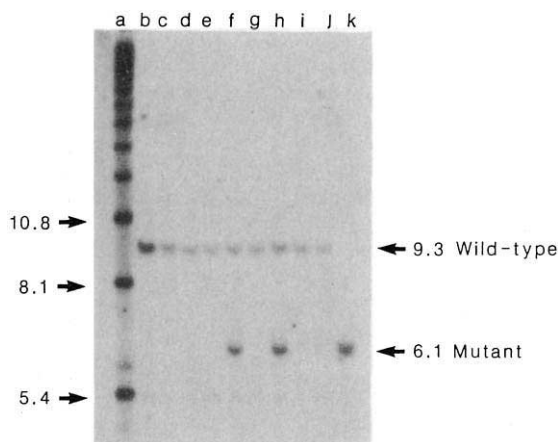


Fig. 2 Southern analysis of DNA from F_1 animals to confirm genetic transmission of the $Hprt-b^{m2}$ allele. DNAs were prepared from tail biopsies taken at approximately two weeks of age, by a modification of the method of Hanahan (from ref. 28). Samples ($\sim 2 \mu\text{g}$) were digested with *Eco*RI and fractionated on a 0.8% agarose gel. Transfer and subsequent steps were as in Fig. 1. The sizes of fragments in the marker lane a are shown beside the lanes. Lane b, wild-type DNA from a female animal; lanes c–e, DNA from male F_1 animals produced by TG3.15, TG4.5 and TG4.13 respectively; lanes f–j, DNA from F_1 females produced by TG4.52, the female germline chimaera; lane k, DNA from the TG R4 clonal line. Two of the F_1 females have the *Eco*RI fragment of ~ 6.1 kb diagnostic for the $Hprt-b^{m2}$ allele, whereas the other three have only the normal pattern and from signal intensity appear to have only one copy of the *Hprt* gene.

gous for the mutant *Hprt* alleles (carrier females) have been bred and approximately equal numbers of male and female F_2 animals obtained (77 females and 66 males). Prenatal lethality due to HPRT deficiency would have led to a 50% reduction in the number of male offspring. The TG.4.52 chimaera has an XO germ line, with the X chromosome carrying the $Hprt-b^{m2}$ allele. This female germline chimaera has produced male progeny which are obligate hemizygous mutants. This provides a further demonstration of the pre- and post-natal viability of mice hemizygous for these mutant *Hprt* alleles.

Following growth in ^3H -hypoxanthine, an autoradiographic analysis of primary fibroblast cultures derived from the hemizygous mutant males produced by TG.4.52 has shown all cells to be completely unlabelled, indicating a total deficiency of HPRT activity (Fig. 3). At eleven weeks of age these males are apparently healthy and remain free from any gross neurological symptoms.

Using HPRT $^-$ mutant stem cells selected in tissue culture we have succeeded in deriving strains of mice carrying mutant *Hprt* alleles in all their cells. Construction of germline chimaeras with stem cells carrying a spontaneous mutation of the *Hprt* gene has been independently achieved by Hooper *et al.*¹⁶. An earlier attempt to derive HPRT $^-$ mice using HPRT $^-$ embryonal carcinoma (EC) cells was not successful, as this clonal line was not capable of colonizing the germ line of chimaeric mice¹⁷.

In human patients with Lesch–Nyhan syndrome there is variation in the level of residual HPRT activity, with the severity of the symptoms usually inversely related to these levels³. One of the most sensitive ways of detecting the low residual levels is through autoradiographic analysis of primary fibroblast cultures^{14,15,18}. Enzyme assays done directly on cell lysates can give completely negative results, but fibroblasts from the same patient will often show labelling¹⁹. To assay HPRT function in animals carrying these newly established alleles, we have applied this test to primary fibroblasts derived from heterozygous female mice and have demonstrated for both alleles that the mutant phenotype originally selected in culture continues to be exhibited in tissues of the mouse. The complete lack of labelling

Table 1 Characteristics of mouse chimaeras

Chimaeras derived from HPRT $^-$ clonal lines

Clonal line	Injected blastocysts	Animals born	Chimaeras		
			Total	Male	Female
TG R3	266	157	66	50	16
TG R4	228	143	94	52	42
TG R6	135	84	43	31	12

Test breeding of chimaeras from HPRT $^-$ clonal lines

Clonal line	Males test bred	Females test bred	Fertile males	Fertile females	Germline chimaeras
TG R3	44	—	36	—	3 males
TG R4	36	25	27	19	8 { 7 males 1 female
TG R6	15	—	6	—	—

Breeding data of germ-line chimaeras

Chimaera	No. of litters	Albino	Pigmented		
			Total	Male	Female
Male TG3.15	2	24	1	1	0
Male TG3.1	6	46	1	0	1
Male TG3.22	6	58	2	1	1
Male TG4.5	3	37	2	1	1
Male TG4.11	2	18	1	0	1
Male TG4.13	3	30	1	1	0
Male TG4.14	2	22	1	1	0
Male TG4.24	1	0	6	4	2
Male TG4.53	2	21	1	1	0
Male TG4.57	1	12	1	1	0
Female TG4.52	3	0	8	2*	6
Total	31	268	25	13	12†

* These two male F_1 animals are hemizygous for *Hprt* mutant alleles.

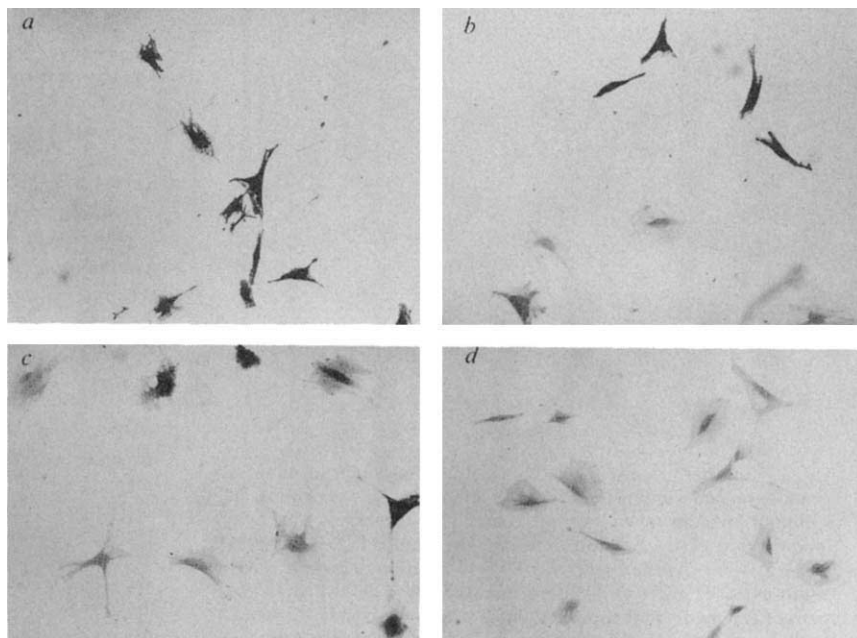
† Female F_1 animals carry *Hprt* mutant alleles.

Three HPRT $^-$ clonal lines were used to produce chimaeras. The chimaeric populations obtained from the TG R3 and TG R6 lines show a sex-ratio distortion effect. The sex ratio of the chimaeras obtained from the TG R4 line was not distorted. Sexually mature chimaeric animals were caged with albino MFI mates and the resulting litters were scored for pigmented progeny. GPI (glucose phosphate isomerase) analysis of blood samples provided confirmatory evidence of the transmission of the EK.CCE-derived *Gpi-1^c* allele. As the TG R4 clonal line is composed of both XO and XY cells both male and female chimaeras were test bred. In the other series only males were set up for test breeding. In all three TG R3 germline chimaeras (bottom part of table) only a minor proportion of functional sperm is derived from the cultured cells, as reflected in the small number of pigmented progeny obtained. The male germline chimaeras from the TG R4 line fall into two categories. Six animals show a partial transmission of the cultured cell genotype, but all offspring from one animal have inherited the dominant coat colour marker. The single female TG R4 germline chimaera has produced a total of 8 offspring (6 females, 2 males) all of which carry the dominant coat-colour phenotype. F_1 progeny were confirmed as carrying the mutant alleles by autoradiographic analysis of cultured fibroblasts and by Southern analysis of DNA obtained from tail biopsies.

Methods. The EK.CCE cell line was derived from a single XY blastocyst-stage embryo of the inbred 129/Sv/Ev strain, homozygous for the *C* (pigmented) and the *Gpi-1^c* alleles, and maintained on STO feeder layers. Before infection the cell line was tested for its ability to form chimaeras by introducing 10–12 individual cells into blastocyst stage embryos as described⁵. Host blastocysts were obtained from MFI outbred strain mice, homozygous for the *c* (albino) and for the *Gpi-1^b* alleles. Live-born animals were sexed at birth and scored later as chimaeric by the presence of coat and eye pigmentation and by GPI analysis. This line has demonstrated 100% chimaera formation and functional germline colonization in 38% of the fertile males. Retroviral infection was designed to produce more than 10 copies per cellular genome in the minimum time. Psi2 cells²⁵ producing the MPSV mos^{-1} neo vector²⁶ at a titre of 10^5 neo R units per ml per 12 h on 3T3 cells were mitotically inactivated by mitomycin treatment and seeded at a density of 10^6 cells on a 6-cm dish. The following day stem cells were plated at a density of 10^6 cells per dish on 10 dishes and cultured on these feeder layers for two days (~ 2 –3 cell doublings). The infected stem cells were collected by trypsinization, mixed and expanded on STO feeder plates. HPRT $^-$ variants were selected by co-plating single-cell suspensions of infected stem cells and mitotically inactivated HPRT $^-$ STO cells directly into medium containing 10^{-6} M 6-TG at a ratio of 2×10^6 stem cells to 5×10^6 STO cells per 10-cm dish on 50 dishes. Colonies growing after two weeks in 6-TG were returned to normal medium, colony-cloned and expanded into clonal cell lines. Chimaeras were made from the TG R3 , 4 and 6 HPRT $^-$ clonal lines by blastocyst injection.

Fig. 3 Autoradiography of cultured primary fibroblasts following growth in medium containing ^3H -hypoxanthine. *a*, Cells derived from XO F_1 female, produced by TG4.52; *b*, cells derived from F_1 female produced by TG3.1; *c*, cells derived from XX F_1 female produced by TG4.52; *d*, cells derived from F_1 male produced by TG4.52.

Methods. Tail biopsies were taken aseptically at birth, minced finely and placed in a drop of medium supplemented with 20% newborn calf serum and non-essential amino acids in a 3-cm dish. A cover-slip was put over the drop and the dish was incubated for several days before collecting fibroblast outgrowths by trypsinization. For the uptake experiment, 10^4 cells were plated on 7-cm Petri-perm culture dishes in 4 ml supplemented medium containing $10 \mu\text{Ci ml}^{-1}$ ^3H -hypoxanthine (8.2 Ci mmol^{-1}). After 24 h, plates were washed twice in phosphate-buffered saline (PBS), fixed in 3:1 methanol/acetic acid vapour, extracted twice with ice-cold 10% trichloroacetic acid (TCA) and washed overnight in running water. Plates were dipped in Kodak NTB-2 emulsion and exposed for 4 days.



of the hemizygous mutant mouse cells in our autoradiographic analysis suggests that there is not even residual HPRT activity in these animals. The DNA polymorphisms associated with each allele provide a convenient molecular marker for identifying animals carrying these mutant *Hprt* genes, obviating the need for the derivation and testing of primary fibroblast cultures.

As is the case in humans, the lack of HPRT activity does not lead to prenatal mortality in mice. The postnatal development and maturation of the male F_1 and F_2 animals carrying the mutant alleles are being followed to determine whether the metabolic, neurological and behavioural symptoms characteristic of the Lesch-Nyhan syndrome are in any way reproduced in an HPRT $^-$ mouse, thereby allowing these animals to serve as models for the human disease. In the human, Lesch-Nyhan males often do not display any behavioural or neurological symptoms until many months after birth³; therefore it may be some time before these mice display any gross abnormalities. Many promising steps towards gene therapy for the Lesch-Nyhan syndrome have been made²⁰⁻²², and HPRT $^-$ mice may provide a useful whole-animal system for advancing this work.

The success of this approach for the genetic manipulation of the mouse opens up the possibility of deriving strains carrying specifically induced alterations in other genes, as long as the mutant cells can be selected from the population of infected cells or easily screened for. It may also eventually be possible to produce specific alterations in endogenous genes through homologous recombination with cloned copies modified *in vitro*^{23,24}.

We thank Wolfram Ostertag and Carol Stocking for the MPSV mos $^{-1}$ neo vector and Drs Thomas Franz, Susan McNabb and Julian Dyson for critical comments. This work was partly supported by a grant from the Cancer Research Campaign. M.R.K. acknowledges a NIH postdoctoral fellowship and the SERC for financial support. A.B. is a Beit Memorial Fellow.

Received 22 January; accepted 13 February 1987.

- Hooper, M. L. *Mammalian Cell Genetics* (Wiley, New York, 1985).
- Littlefield, J. W. *Science* **145**, 709-710 (1964).
- Melton, D. W., Konecki, D. S., Brennand, J. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2147-2151 (1984).
- Nesbitt, M. N. & Franke, U. *Chromosoma* **41**, 145-158 (1973).
- Burgoyne, P. S. & Baker, T. G. *J. Reprod. Fert.* **61**, 207-212 (1981).
- Rosenbloom, F. M., Kelly, W. N., Henderson, J. F. & Seegmiller, J. E. *Lancet* **ii**, 305-306 (1967).
- Migeon, B. R., Der Kaloustian, V. K., Nyhan, W. L., Young, W. J. & Childs, B. *Science* **160**, 425-427 (1968).
- Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292-295 (1987).
- Dewey, M. J., Martin, D. W. Jr, Martin, G. R. & Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5564-5568 (1977).
- Migeon, B. R. *Biochem. Genet.* **4**, 377-383 (1970).
- Fujimoto, W.-Y. & Seegmiller, J. E. *Proc. natn. Acad. Sci. U.S.A.* **65**, 577-584 (1970).
- Willis, R. C. *et al. J. biol. Chem.* **259**, 7842-7849 (1984).
- Miller, A. D., Eckner, R. J., Jolly, D. J., Friedmann, T. & Verma, I. M. *Science* **225**, 630-632 (1984).
- Patel, P. I. & Caskey, C. T. *BioEssays* **2**, 4-8 (1986).
- Thomas, K. R., Folger, K. R. & Cappechi, M. R. *Cell* **44**, 419-428 (1986).
- Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. *Nature* **317**, 230-234 (1985).
- Mann, R., Mulligan, R. C. & Baltimore, D. *Cell* **33**, 153-157 (1983).
- Stocking, C., Kollek, R., Bergholz, U. & Ostertag, W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5746-5750 (1985).
- Konecki, D. S., Brennand, J., Fuscoe, J. C., Caskey, C. T. & Chinault, A. C. *Nucleic Acids Res.* **10**, 6763-6775 (1982).
- Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo—a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1986).

Alternative pathway activation of T cells by binding of CD2 to its cell-surface ligand

T. Hünig*, G. Tiefenthaler*,
K.-H. Meyer zum Büschenfelde† & S. C. Meuer†

* Genzentrum der Universität München, Am Klopferspitz, D-8033 Martinsried, FRG

† I. Medizinische Klinik und Poliklinik, Johann Gutenberg Universität, Langenbeckstr. 1, D-6500 Mainz, FRG

Activation of resting T lymphocytes is initiated by the interaction of cell-surface receptors with their corresponding ligands. In addition to activation through the CD3 (T3)-Ti antigen-receptor complex¹, recent experiments have demonstrated induction of T-cell proliferation through the CD2 (T11) molecule²⁻⁴, traditionally known as the erythrocyte(E)-receptor, through which T cells can bind red blood cells (RBC)⁵⁻⁷. This 'alternative pathway' of T-cell

- Lesch, M. & Nyhan, W. L. *Am. J. Med.* **36**, 561-570 (1964).
- Seegmiller, J. E., Rosenbloom, F. M. & Kelley, W. N. *Science* **155**, 1682-1684 (1967).
- Kelley, W. N. & Wyngaarden, J. B. in *The Metabolic Basis of Inherited Diseases* 5th edn (eds Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L. & Brown, M. S.) 1115-1143 (McGraw-Hill, New York, 1983).
- Stout, J. T. & Caskey, C. T. *A. Rev. Genet.* **19**, 127-148 (1985).
- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154-156 (1981).
- Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. *Nature* **309**, 255-256 (1984).
- Robertson, E., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445-448 (1986).
- King, W., Patel, M. D., Lobel, L. I., Goff, S. P. & Nguyen-Huu, C. *Science* **228**, 554-558 (1985).

activation² was observed *in vitro* in response to combinations of anti-CD2 monoclonal antibodies (mAbs) that bind to distinct epitopes of CD2, such as mAbs against T11₂ plus T11₃ (ref. 2). The physiological importance of this activation pathway can be assessed only by studying the effects of a naturally occurring ligand of CD2 on T-cell activation. We have recently described such a ligand, a glycoprotein of apparent relative molecular mass 42,000 (M_r 42K) that is expressed on all blood cells and some other tissues^{9,11}. Here we demonstrate that binding of this cell surface molecule, termed T11 target structure or T11TS, to CD2 (T11) induces reactivity in resting T cells to a mitogenic stimulus given by a mAb to the T11₃ determinant or by submitogenic concentrations of anti-T11₂₊₃ mAbs. Thus, one of the signals required for T-cell activation through the alternative pathway is provided by the interaction of CD2 with a naturally occurring complementary cell-surface molecule.

Our experimental approach was based on the hypothesis that CD2 is a cell-interaction molecule with a complementary ligand, T11TS, that is expressed not only on RBC, where it is involved in E-rosetting⁸, but also on other cell types, where it plays a role in T-cell activation⁹. CD2 binds to both human and sheep RBC^{8,10}, so it apparently recognizes its ligand across this species barrier. We have identified sheep T11TS with a mAb to sheep RBC (SRBC) that completely blocks their binding to human or sheep T cells⁹. This mAb, called L180/1, detects T11TS on a variety of cell types, including most if not all blood cells^{9,11}. Its functional involvement in T-cell activation was suggested by its ability to block the sheep mixed lymphocyte reaction⁹. Although mAb L180/1 recognizes a sheep-specific determinant on T11TS⁸, recent evidence has indicated that the lymphocyte function associated antigen (LFA)-3¹² may represent the human form of T11TS. Thus, anti-LFA-3 mAb blocks binding of human RBC to T cells²¹, and the LFA-3 antigen has a cellular distribution similar to T11TS¹². Moreover, a polyclonal rabbit anti-sheep T11TS antiserum inhibits human autologous E-rosetting (G.T. and T.H., in preparation). Evidence that T11TS and LFA-3 share antigenic determinants is shown in Fig. 1. Human RBC were stained with a mAb to LFA-3 (from Dr T. Springer) in the presence or absence of rabbit anti-T11TS antiserum or preimmune serum. Flow cytometric analysis revealed that the anti-T11TS antiserum, but not the preimmune serum, completely blocks the binding of the anti-LFA-3 mAb to human RBC, indicating that it contains antibodies that mask the LFA-3 determinant. Reactivity of mAbs to human RBC unrelated to the LFA-3 molecule was not affected by the antiserum (not shown). Taken together with the functional data mentioned above, this is a strong indication that T11TS and LFA-3 are both functional and structural homologues in sheep and man, respectively.

The presence of T11TS/LFA-3 on all types of blood cells, including the responding T cells, makes it difficult to pinpoint the site of action of anti-T11TS or anti-LFA-3 antibodies in T-cell activation studies. This problem can be overcome, however, if human T cells are used as responder cells and SRBC as a source of T11TS-expressing cells together with a sheep-specific anti-T11TS mAb. Accordingly, we have employed SRBC, mAb L180/1 and the anti-T11₂ and anti-T11₃ mAbs to investigate whether binding of a naturally occurring ligand to CD2 provides an activation signal to human T lymphocytes. Nylon-wool enriched T lymphocytes from peripheral blood were cultured with or without SRBC in the presence of either a combination of anti-T11₂ and anti-T11₃ mAbs, or anti-T11₃ alone to see if SRBC could replace the requirement for the T11₂-specific mAb (Fig. 2). This approach was chosen because anti-T11₂ or anti-T11₃ antibodies by themselves are not mitogenic, but together activate T lymphocytes polyclonally². If binding of T11TS to CD2 is one of the steps physiologically involved in this 'alternative pathway', such a signal may be expected to replace the requirement for one of the two anti-T11 mAbs in T-cell activation through the CD2. As expected, a proliferative response to anti-T11₂ plus anti-T11₃, but not to either antibody

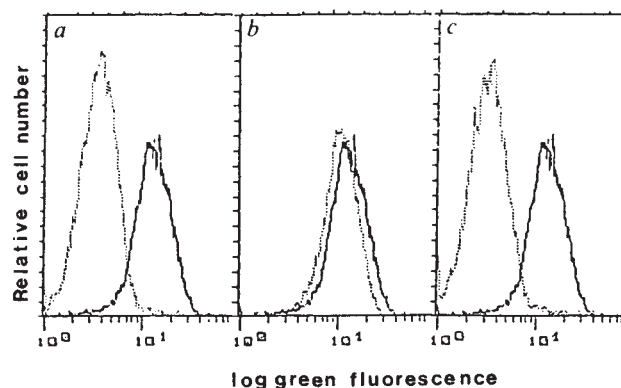


Fig. 1 Rabbit anti-T11TS antiserum blocks binding of anti-LFA-3 mAb to human RBC. *a*, No rabbit serum. Straight line, anti-LFA-3; dotted line, negative control. *b*, Anti-LFA-3 in the absence (straight line) or presence (dotted line) of preimmune serum. *c*, Anti-LFA-3 in the absence (straight line) or presence (dotted line) of rabbit anti-T11TS serum.

Methods. Human RBC (5×10^5) were treated with the mouse mAb TS2/9¹² against LFA-3 at $1 \mu\text{g ml}^{-1}$ or with a negative control mAb of the same isotype in the presence or absence of rabbit preimmune serum or anti-T11TS immune serum (1:10 final dilution). After 30 min on ice, the cells were washed, stained for 30 min with fluorescein isothiocyanate (FITC)-conjugated Fab₂ rabbit anti-mouse immunoglobulin (DAKO, 1:50), and analysed on a FACScan flow cytometer (Becton-Dickinson) using a light scatter gate for individual RBC and logarithmic amplification for fluorescence. Rabbit anti-T11TS antiserum was prepared by injecting $50 \mu\text{g}$ aluminium hydroxide-adsorbed sheep T11TS purified to homogeneity by immune-affinity and preparative SDS-polyacrylamide gel electrophoresis¹¹, followed by five injections of $50 \mu\text{g}$ T11TS in soluble form at 3-week intervals.

by itself, was observed in the absence of SRBC (Fig. 2, experiment 1). In contrast, anti-T11₃ alone could induce proliferation if SRBC were present. Addition of either intact mAb L180/1 or of its Fab fragment completely abolished the stimulatory effect of SRBC, indicating that binding of T11TS to CD2 was essential. In addition, a vigorous proliferative response to a submitogenic concentration of T11₂ plus T11₃-specific mAbs was observed in the presence of SRBC (Fig. 2, experiment 2). Again, this response was abolished by mAb L180/1. It is unlikely that this inhibition is due to a non-specific effect exerted by, for example, antibody-coated SRBC because a different SRBC-specific mAb of the same isotype and used at the same titre as mAb L180/1 was unable to reverse the co-stimulation by SRBC (Fig. 3). Thus, mAb L180/1 appears to neutralize the stimulatory effect of T11TS. Figure 3 also demonstrates that T11TS-dependent induction of the alternative pathway is a dose-dependent phenomenon: the response to SRBC plus anti-T11₃ increases steeply with the addition of greater numbers of SRBC. Again, this effect is completely blocked by the addition of anti-T11TS mAb.

These results show that the interaction of CD2 with its naturally occurring complementary cell-surface molecule provides an activation signal to T lymphocytes through the 'alternative pathway'. At first sight, it appears paradoxical that despite the expression of T11TS (or LFA-3) on leukocytes, including T cells, SRBC are required to demonstrate the functional effects of CD2-T11TS interaction. The inability of T-lymphocyte-borne T11TS to complement anti-T11₃ or submitogenic concentrations of anti-T11₂₊₃ in T-cell activation is most probably due to the high negative net charge of resting peripheral T cells¹³ that prevents physical contact between them¹⁴. This interpretation is supported by the fact that despite the expression of LFA-3 on human RBC, the formation of autologous E-rosettes does not occur *in vivo* but requires reduction of negative surface charge of either the lymphocytes or the RBC¹⁵ by neuraminidase

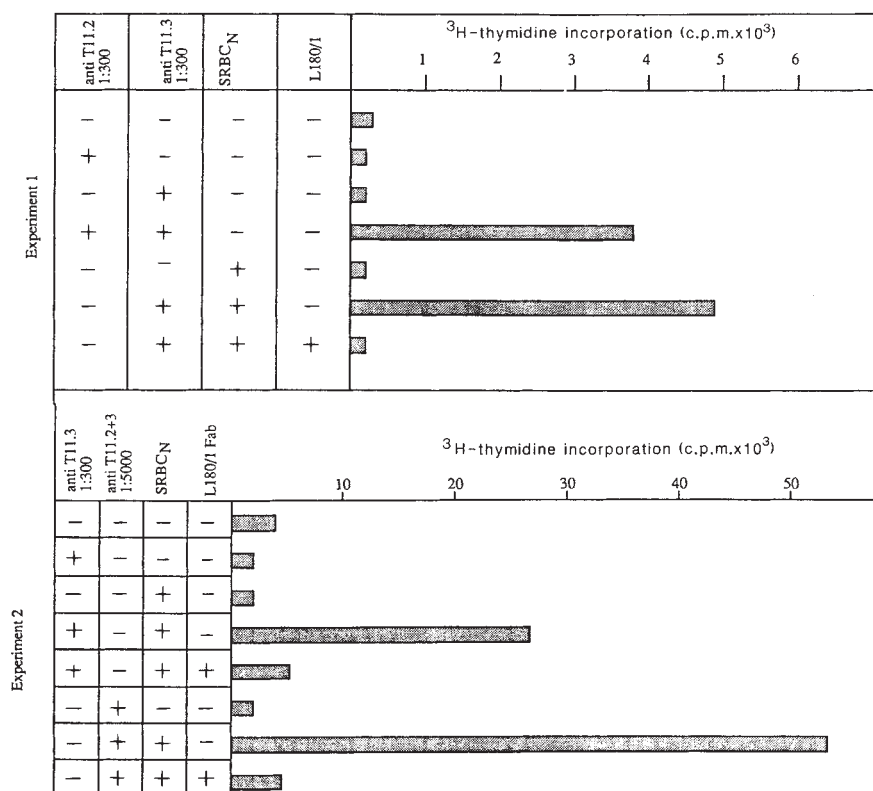


Fig. 2 Co-stimulatory effect of SRBC is inhibited by anti-T11TS mAb. Peripheral blood mononuclear cells (PBMNC) were isolated by Ficoll density centrifugation and passed over nylon wool to enrich for T lymphocytes. Cells (10^5) were cultured for 4 days in round-bottom microwells in a final vol. of 0.2 ml RPMI 1640 medium containing 10% heat-inactivated human serum, antibiotics and 2% glutamine, with the additions indicated. Ascites fluids containing anti-T11₂ and anti-T11₃ mAbs were used at the concentrations given. L180/1 was purified antibody used at $10 \mu\text{g ml}^{-1}$. Antigen-binding fragment (Fab) of L180/1 ($10 \mu\text{g ml}^{-1}$) was prepared by papain digestion and shown to contain no intact antibody. SRBC (10^6 per well) were treated with neuraminidase (to give SRBC_N) to reduce surface charge and promote E-rosetting. At the onset of culture, the plates were centrifuged for 5 min at 200g to facilitate binding of SRBC to T lymphocytes. Cultures were pulsed with $1 \mu\text{Ci } ^3\text{H}$ -thymidine for the last 16 h of culture. Standard deviations of results from triplicate cultures were below 15%.

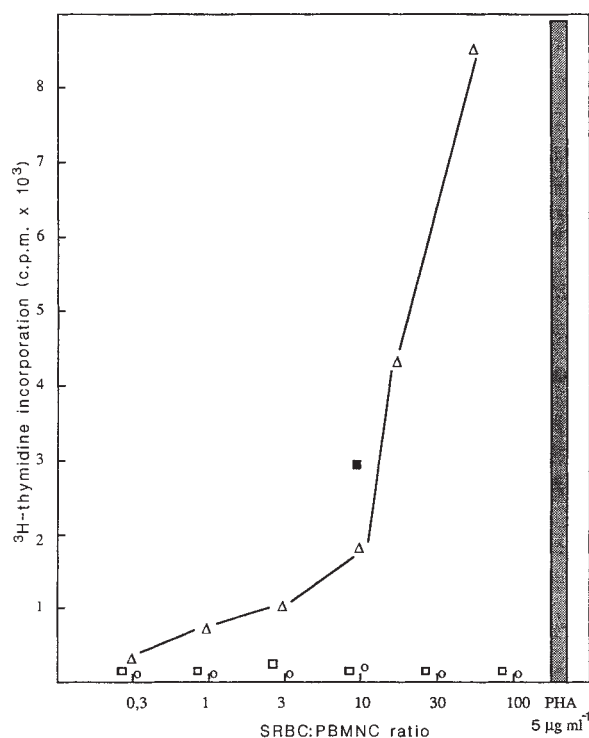


Fig. 3 Dose-response relationship of SRBC-induced T-cell activation by the alternative pathway. Responses are shown to: $\square$, SRBC alone; Δ , SRBC + anti-T11₃; $\circ$, SRBC + anti-T11₃ + L180/1; $\blacksquare$, SRBC + anti-T11₃ + 1/2. Culture conditions were as in Fig. 2 except that SRBC_N were added at 3×10^4 to 1×10^7 per well, as indicated. Anti-T11₃ was used at 1:300 dilution of ascitic fluid. Like L180/1, mAb 1/2 is an immunoglobulin G (IgG)₁ antibody reactive with both SRBC and sheep leukocytes, but is not specific for T11TS. Also shown is the response to $5 \mu\text{g ml}^{-1}$ phytohaemagglutinin (PHA, Sigma). PBMNC, peripheral blood mononuclear cells.

treatment¹⁶. SRBC bind readily to resting T cells through CD2-T11TS and thus provide a system to study the functional effects of this interaction at the level of the resting T cells. We suggest that the function performed by SRBC in this experimental situation is performed *in vivo* by other cell types to which T lymphocytes or thymocytes bind through CD2, such as activated antigen-presenting cells or thymic epithelial cells¹⁷.

Despite the clear-cut functional effect of CD2-T11TS interaction, we are far from understanding the whole sequence of events involved in T-cell activation through CD2. In particular, the natural counterpart to the 'second signal' required in addition to T11TS (here given by anti-T11₃ or submitogenic concentrations of anti-T11₂₊₃) remains unresolved. A requirement for two signals in T-cell activation through CD2 was also apparent in the recent experiments of Holter *et al.*⁴, who confirmed that anti-CD2 mAbs induce T-cell proliferation only if added in pairs. Interestingly, they also found that in the presence of the protein kinase C activator phorbol myristate acetate, individual anti-CD2 mAbs can also activate T lymphocytes. The latter finding suggests that under physiological conditions, the interaction of CD2 with T11TS could also be complemented by a CD2-unrelated activation pathway. Sub-optimal signalling through the antigen-receptor complex has recently been shown to synergize with an individual anti-CD2 mAb in T-cell activation¹⁸. Further support for this view comes from the recent finding that binding of CD2 to the human LFA-3 molecule in cytotoxic T lymphocyte (CTL)-target cell conjugation is also by itself insufficient to trigger cytolysis¹⁹. One possibility is that an additional signal required for T-cell activation through CD2 is a lymphokine. The results of Larsson *et al.*²⁰, who before the discovery of the CD2 described acquisition of reactivity to 'mitogenic factors' as a consequence of E-rosetting, are compatible with such a mechanism. In conclusion, the interaction of CD2 with the naturally occurring ligand T11TS provides an activation signal to T lymphocytes which by itself is insufficient to trigger T-cell activation. How this step fits into the series of events that are part of physiologically occurring T-cell responses is currently under investigation.

We thank Dr T. Springer for mAb TS2/9, Dr E.-L. Winnacker for continuous support, and Ms M. Haurer for technical assistance. This work was supported by grants from the Bundesministerium für Forschung und Technologie, DFG-SFB 311/A3, and by Genzentrum München e.V.

Received 10 October 1986; accepted 7 January 1987.

- Meuer, S. C. *et al.* *Nature* **303**, 808–810 (1983).
- Meuer, S. C. *et al.* *Cell* **36**, 897–906 (1984).
- Brottier, P., Boumsell, B., Gelin, C. & Bernard, A. *J. Immun.* **135**, 1624–1631 (1985).
- Holter, W., Fischer, F., Majdic, O., Stockinger, H. & Knapp, W. *J. exp. Med.* **163**, 654–664 (1986).
- Wortis, H. H., Cooper, A. C. & Brown, M. C. *Nature new Biol.* **243**, 109–111 (1973).
- Howard, F. D. *et al.* *J. Immun.* **126**, 2117–2122 (1981).
- Kamoun, M., Kadin, M. E., Martin, P. J., Nettleton, J. & Hansen, J. A. *J. exp. Med.* **153**, 207–212 (1981).
- Hünig, T. *J. exp. Med.* **162**, 890–901 (1985).
- Hünig, T. *J. Immun.* **136**, 2103–2108 (1986).
- Scheffel, J. W. & Swartz, S. J. *J. Immun.* **128**, 1930–1932.
- Hünig, T., Mitnacht, R., Tiefenthaler, G., Köhler, C. & Miyasaka, M. *Eur. J. Immun.* (in the press).
- Krensky, A. M. *et al.* *J. Immun.* **131**, 611–616 (1983).
- Bert, G., Forrester, J. A. & Davies, A. J. S. *Nature new Biol.* **234**, 86–87 (1971).
- Hünig, T. *Eur. J. Immun.* **13**, 596–601 (1983).
- Seiler, F. R., Sedlacek, H. H., Kanzy, E. J. & Lang, W. *Behring Inst. Mitt* **52**, 26–72 (1972).
- Baxley, G., Bishop, G. B. & Wortis, H. H. *Clin. exp. Immun.* **15**, 385–392 (1973).
- Singer, K. H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 6588–6592 (1986).
- Yang, S. Y., Chouaib, S. & Dupont, B. *J. Immun.* **137**, 1097–1100 (1986).
- Shaw, S. *et al.* *Nature* **323**, 262–264.
- Larsson, E.-L., Andersson, J. & Coutinho, A. *Eur. J. Immun.* **8**, 693–696 (1978).
- Plunkett, M. L., Sanders, M. E., Selvaraj, P., Dustin, M. L. & Springer, T. A. *J. exp. Med.* (in the press).

Ion channels activated by inositol 1,4,5-trisphosphate in plasma membrane of human T-lymphocytes

Miyuki Kuno* & Phyllis Gardner†

Department of Medicine, Stanford University School of Medicine, Stanford, California 94305, USA

Hydrolysis of membrane-associated phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)-P₂) to water soluble inositol 1,4,5-trisphosphate (Ins(1,4,5)P₃) is a common response by many different kinds of cells to a wide variety of external stimuli (see refs 1 and 2 for review). Ins(1,4,5)P₃ is a putative second messenger which increases intracellular Ca²⁺ by mobilizing internal Ca²⁺ stores, a hypothesis which has been substantiated by studies with chemically permeabilized cells^{3–13} and with isolated microsomal membrane fractions^{11,12–16}. But the possibility that Ins(1,4,5)P₃ could induce in intact cells an influx of external Ca²⁺ through transmembrane channels, originally hypothesized by Michell in 1975¹⁷, has never been directly tested. We report here single-channel recordings of an Ins(1,4,5)P₃-activated conductance in excised patches of T-lymphocyte plasma membrane. The Ins(1,4,5)P₃-activated transmembrane channel appears to be identical to the recently described mitogen-regulated, voltage-insensitive Ca²⁺ permeable channel involved in T-cell activation¹⁸. We suggest that Ins(1,4,5)P₃ acts as the second messenger mediating transmembrane Ca²⁺ influx through specific Ca²⁺-permeable channels in mitogen-stimulated T-cell activation.

Single-channel currents were recorded from excised inside-out patches of the Jurkat E6-1 human T-cell line¹⁹ by the extracellular patch clamp technique²⁰. In general, recording pipettes contained isotonic, tetrodotoxin (TTX)-supplemented BaCl₂, though similar results were obtained with 'physiological' external solution ([Ca²⁺]_o = 2 mM). Ba²⁺ was chosen as the current carrier over Ca²⁺ because Ba²⁺ blocks the K⁺ currents present in T-cell membranes^{21–23}, and because Ba²⁺ has been shown to be more permeant, and thus improve signal resolution through traditional Ca²⁺ channels^{24–27}. In all cases, the bath contained

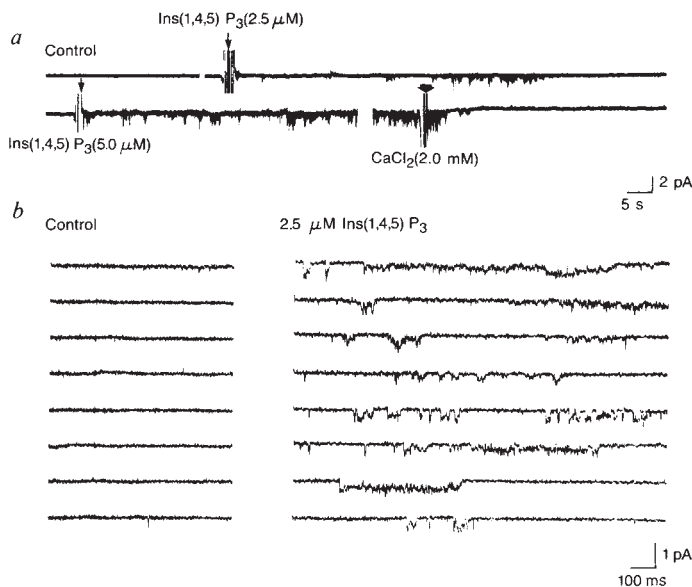


Fig. 1 Single Ins(1,4,5)P₃-induced inward currents recorded from an inside-out membrane patch excised from a Jurkat cell. The patch pipette contained 110 mM BaCl₂, 200 nM TTX, and 10 mM HEPES-KOH (pH 7.3). The membrane patch cytoplasmic surface was bathed in a solution containing: 150 mM KCl, 0.55 mM CaCl₂, 2 mM MgCl₂, 1.1 mM EGTA, and 10 mM HEPES-KOH (pH 7.3). The free internal Ca²⁺ concentration was estimated at 10⁻⁷ M using an apparent dissociation constant of 10⁻⁷ (ref. 48). The holding potential was 0 mV. Inward currents are seen as downward deflections from baseline. All records were filtered at 1 kHz (8-pole Bessel). *a*, Ins(1,4,5)P₃-induced inward currents from one excised inside-out patch at slow sweep speed. Arrows, addition of Ins(1,4,5)P₃ (Sigma) and CaCl₂ (final concentrations in parentheses) to the bath. The ringing transients during applications are secondary to vibration. No inward currents were observed in the patch for >5 min before addition of Ins(1,4,5)P₃. The first break in the recording indicates an interruption of 4 min during which the potential was changed. The first bath application of Ins(1,4,5)P₃ resulted in inward current activity within 10 s of addition to the bath. Inward currents appeared <60 s and then disappeared for a period of >4 min. Upper and lower traces were interrupted during this interval. The second application of Ins(1,4,5)P₃ (final concentration 5.0 μM) immediately elicited inward currents which were sustained for longer than 3 min. The interruption in the trace indicates a 2-min interval during which the membrane potential was changed. The inward currents were abolished by the application of CaCl₂ (final concentration 2.0 mM) to the internal bathing solution. Current amplitudes were slightly attenuated owing to the limited bandwidth of the pen recorder (Gould 2400S). *b*, Ins(1,4,5)P₃-induced unitary inward currents at higher time resolution with a sampling rate of 4 kHz. Left sequential traces, control recordings before the applications of Ins(1,4,5)P₃. Right traces, inward current traces in the presence of 2.5 μM Ins(1,4,5)P₃. Right traces were not consecutive in time, but selected to illustrate features of inward currents associated with Ins(1,4,5)P₃ (see text). **Methods.** All experiments were performed on cells of the malignant T-lymphocyte cell line Jurkat E6-1, which were maintained in medium RPMI-1640, supplemented with 10% heat-inactivated fetal calf serum, 50 μg ml⁻¹ gentamycin, 2 mM L-glutamine (Gibco), at 37°C in humidified 5% CO₂ atmosphere. Before recording, the cells were washed twice and resuspended in the bathing solution listed above. All experiments were performed at room temperature (20–22°C). Single-channel currents were recorded using the cell-attached and inside-out recording configurations²⁰ with a List-EPC7 patch-clamp amplifier. The Sylgaard-coated patch pipettes were pulled from thin-walled borosilicate glass pipettes to a lumen diameter of ≤0.5 μm. Current signals were stored on analogue tape and digitized for analysis.

high KCl (150 mM), low Ca²⁺ (10⁻⁷ M) solution to mimic inter-nal ionic composition.

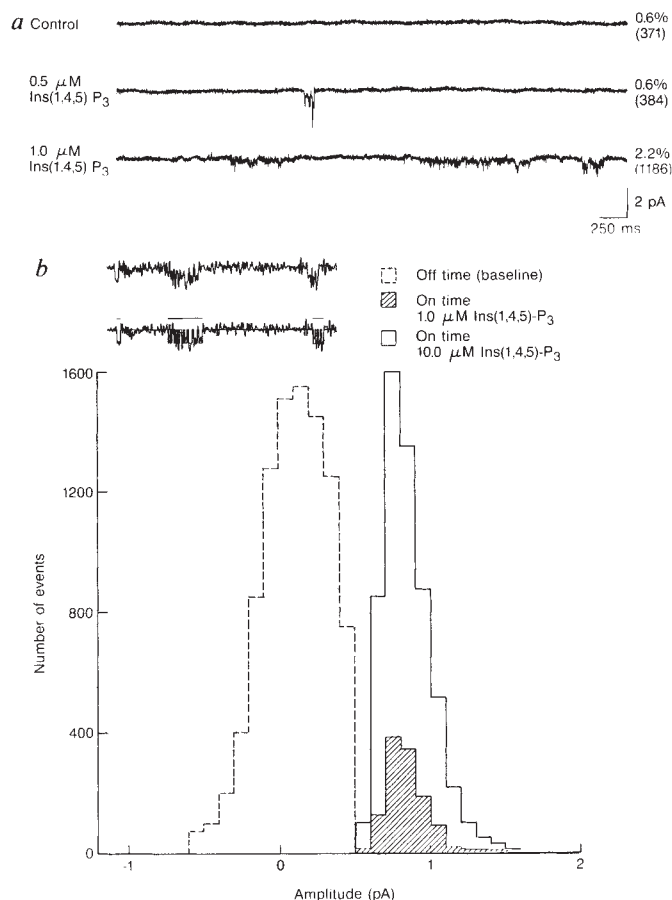
Figure 1a shows the current recorded from one excised patch held at 0 mV. Within seconds after addition of Ins(1,4,5)P₃ (final

* Present address: Department of Physiology, Osaka City University Medical School, Abenoku, Osaka 545, Japan.

† To whom correspondence should be addressed.

Fig. 2 Effect of $\text{Ins}(1,4,5)\text{P}_3$ concentration on single channel inward current activity. *a*, Inward current traces from the same inside-out patch in control internal solution and internal solution supplemented with 0.5 or 1.0 μM $\text{Ins}(1,4,5)\text{P}_3$ as indicated. Numbers at right indicate percent open-time (total open-time over total duration of record) and total number of openings, in parentheses, in 62.259-second continuous recordings under each condition. Recording conditions are as in Fig. 1. *b*, Probability density histogram of current amplitude in the presence of 1.0 μM $\text{Ins}(1,4,5)\text{P}_3$ (cross-hatched area) and 10.0 μM $\text{Ins}(1,4,5)\text{P}_3$ (continuous profile). Recording at top left is representative digitized data before and after superposition of computer-derived idealized trace of baseline and singles. (Lines above trace indicate computer-derived bursts, see for example Fig. 3c). The dashed profile indicates the off-time (baseline, corresponding to 0 pA) distribution in the presence of 10.0 μM $\text{Ins}(1,4,5)\text{P}_3$. The off-time distribution in the presence of 1.0 μM $\text{Ins}(1,4,5)\text{P}_3$ was similar, but with smaller area, as the number of off events is simply $(n-1)$ open events. For both 1.0 and 10.0 μM concentrations, on-time distributions were fitted from 0.50 pA, the lower single-event detection threshold, which was approximately 2.5 times the mean deviation of baseline noise (computer-derived determination of 0.2068 pA). Period for analysis was 32.768 s. The total number of detected events were 652 for 1.0 μM and 2365 for 10.0 μM . The peak of the on-time amplitude histograms corresponded to a unitary amplitude of 0.71 ± 0.11 pA for 1.0 μM $\text{Ins}(1,4,5)\text{P}_3$ and 0.71 ± 0.14 pA for 10.0 μM $\text{Ins}(1,4,5)\text{P}_3$. The area ratio under the two on-time amplitude histograms was 5.0 (5,597 events-pA: 1,117 events-pA), indicating a greater number of events at the higher concentration.

Methods. For data analysis, records on analogue tape were digitized at 10 kHz and stored on a microcomputer (IBM PC-AT). The automated single-channel analysis program of Sachs *et al.*⁴⁹ implemented by C. Lingle (personal communication), was used to produce the plots shown in *b* and in Fig. 3b and c and to calculate kinetics in Table 1. The baseline level (corresponding to 0 pA) was located using an algorithm that maximizes the number of crossings through a horizontal line. An event detection threshold was set at 0.5 times the predetermined single channel amplitude (by eye, which was slightly higher than the computer-derived mean owing to inclusion of short events which do not reach full amplitude), which was confirmed to be 2.5–3.0 times the computer-derived mean baseline noise in each case. Times of transition were defined as crossings of the midline between the baseline and open current level. Multiple event exclusion threshold was set at 1.8 times the predetermined single-channel amplitude. Interval analysis was used to generate distributions of open times, closed times, and burst durations (as in Fig. 3), which were fitted with probability density functions by the method of maximum likelihood to the continuous distribution of data. The first five bins (500 μs) of the open-time and short closed-time distributions were excluded from curve-fitting because of error arising from events inadequately resolved owing to bandwidth attenuation. Patches used for kinetic analysis were selected to contain few, if any, multiple events.



concentration 2.5 μM) to the bath, inward currents appeared. At low concentrations (1.0–2.5 μM), the effect of $\text{Ins}(1,4,5)\text{P}_3$ appeared to be transient, usually lasting 10–60 s. A second addition of $\text{Ins}(1,4,5)\text{P}_3$ (final concentration 5.0 μM) promptly elicited sustained inward current pulses. The inward currents abruptly disappeared after addition of millimolar concentrations of Ca^{2+} to the bath, indicating that the currents were not secondary to patch membrane deterioration and suggesting a possible cytosolic Ca^{2+} inactivation process. Figure 1b shows the single-channel inward current recorded from another patch at higher time resolution. Addition of 2.5 μM $\text{Ins}(1,4,5)\text{P}_3$ elicited inward currents, which consisted of intermittent bursts of brief openings, separated by quiet intervals. Unitary current amplitude was ~ 0.6 pA at 0 mV and slope conductance approximately 7 pS (derived from single-channel $I-V$ relationship, not shown).

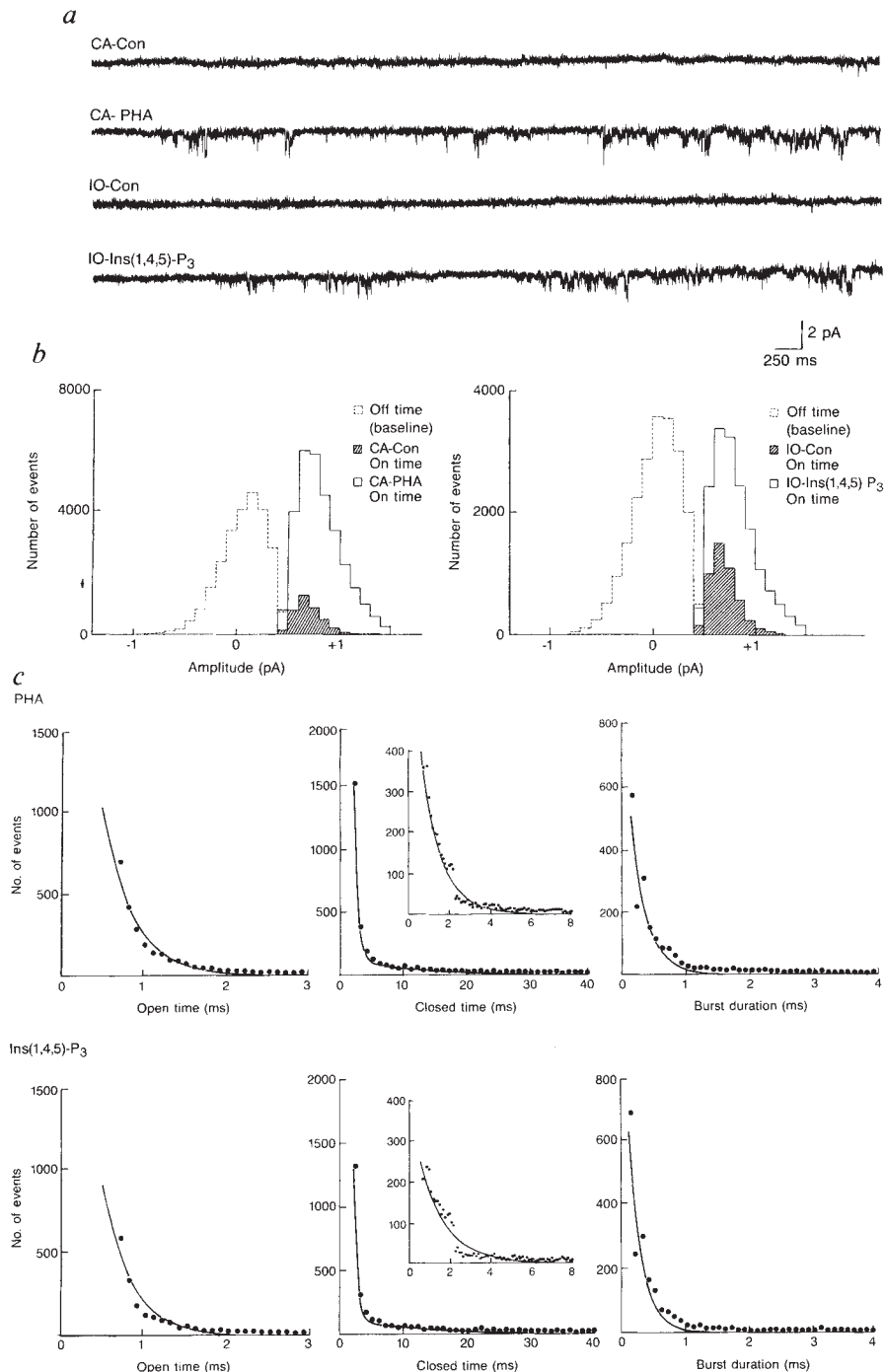
Similar results were obtained in 14 patches. In 3 of these, inward currents were observed only after the addition of $\text{Ins}(1,4,5)\text{P}_3$ to the bath, despite observation periods of longer than 5 min for the control period. In 11 other patches, $\text{Ins}(1,4,5)\text{P}_3$ greatly augmented the inward currents. In all cases, subsequent addition of micromolar or greater Ca^{2+} concentrations to the bath appeared to inactivate the $\text{Ins}(1,4,5)\text{P}_3$ -activated inward currents. Evidence that suggested that the $\text{Ins}(1,4,5)\text{P}_3$ -activated inward current flowed through Ca^{2+} -permeable channels included the following: (1) Ba^{2+} was the only charge carrier with an inwardly directed driving force at 0 mV in the TTX-containing, Na^+ -free pipette solution; (2) The amplitude of the single-channel current decreased with increasing depolarization

from 0 mV, with no clear-cut reversal identified, indicating a positive reversal potential E_{rev} .

Estimates of the concentration that $\text{Ins}(1,4,5)\text{P}_3$ can reach in stimulated cells, obtained from direct measurement of labelled $\text{Ins}(1,4,5)\text{P}_3$, fall in the range 3–75 μM in a variety of cells^{3,28–31}. The dependence of channel activation on micromolar $\text{Ins}(1,4,5)\text{P}_3$ concentrations, which are in the apparent physiological range, was investigated. Activation, observed at threshold concentrations between 0.5 and 1.0 μM (Fig. 2a), was dose-dependent at least up to 10 μM . As concentration increased from 1.0 μM to higher concentrations, increasing augmentation of inward current was seen. Augmentation did not occur through a change in unitary current amplitude, but through an apparent increase in frequency of single-channel events. Figure 2 shows the amplitude histogram from one patch measured for 33 s after applications of 1.0 μM and 10.0 μM $\text{Ins}(1,4,5)\text{P}_3$, respectively. A greater number of opening events were detected at the higher concentration ($n = 652$ for 1.0 μM and $n = 2365$ for 10.0 μM) and the area under the probability distribution curve was correspondingly five times greater, but mean amplitudes of unitary current (0.71 ± 0.11 pA for 1.0 μM and 0.71 ± 0.14 pA for 10.0 μM , mean $\pm$ s.d.) were virtually identical. From similar examples, we conclude that $\text{Ins}(1,4,5)\text{P}_3$ augments inward Ba^{2+} currents with increasing concentrations at least in part by increasing the probability of channel opening.

We have recently demonstrated a voltage-insensitive, Ca^{2+} -permeable channel in the membrane of T lymphocytes¹⁸ which has an increased probability of opening following stimulation

Fig. 3 Comparison of PHA- and $\text{Ins}(1,4,5)\text{P}_3$ -induced single channel currents. Recordings were made in the cell-attached and inside-out configuration. Pipette solution contained 110 mM BaCl_2 , 200 nM TTX, and 10 mM HEPES-KOH (pH 7.3). Bath solution contained 150 mM KCl, 1 mM MgCl_2 , 0.55 mM CaCl_2 , 1.1 mM EGTA, and 10 mM HEPES-KOH (pH 7.2). The holding potential was 0 mV in all recordings (cellular resting potential was assumed to be zeroed by the isotonic KCl bath solution). Current signals were filtered at 1 kHz. All records are from the same patch. Probability density histograms were constructed from recordings of 62.259 s for each condition. **a**, Inward currents from a cell-attached patch before (CA-Con) and 5 min after bath application of $6 \mu\text{g ml}^{-1}$ PHA (CA-PHA), and from the patch isolated into the inside-out configuration before (IO-Con), and 20 s after bath application of $2.5 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$ (IO- $\text{Ins}(1,4,5)\text{P}_3$). Inward currents are seen as downward deflections from baseline. **b**, Probability density histograms of on-time current amplitudes from the patch in the cell-attached configuration before (cross-hatched area) and after (continuous profile) PHA application are in the left panel. On-time amplitude histograms are fitted from 0.4 pA, the lower limit of single event detection threshold, which was >3.0 times the computer-derived mean deviation of baseline noise of 0.13 pA. Dashed line, off-time amplitude (baseline) distribution in the CA-PHA recording. The peak of the on-time amplitude histograms corresponded to a unitary amplitude of 0.57 ± 0.12 pA in the control cell-attached recording and 0.65 ± 0.19 pA after PHA application. The area ratio of PHA to control on-time distributions was 7.9 (28,660 events-pA: 3,620 events-pA). Probability density histograms of on-time current amplitudes from the same patch after isolation into the inside-out configuration before (cross-hatched area) and after (continuous profile) addition of $2.5 \mu\text{M}$ $\text{Ins}(1,4,5)\text{P}_3$ to the bath solution are in the right panel. The dashed line represents the off-time amplitude (baseline) distribution in the IO- $\text{Ins}(1,4,5)\text{P}_3$ inside-out recording. The peak of the on-time amplitude histograms corresponded to a unitary amplitude of 0.57 ± 0.12 pA in the control inside-out patch recording, and 0.63 ± 0.18 pA after bath application of $\text{Ins}(1,4,5)\text{P}_3$. The area ratio of $\text{Ins}(1,4,5)\text{P}_3$ to control was 3.5 (15,530 events-pA: 4,456 events-pA). **c**, Distributions of open times, closed times, and burst durations for the PHA-induced (upper panels) and $\text{Ins}(1,4,5)\text{P}_3$ -activated (lower panels) inward currents. The durations of the two records analysed were 62.259 s, with a total of 5,689 openings (CA-PHA) and 4,659 openings (IO- $\text{Ins}(1,4,5)\text{P}_3$). The open-time and burst duration histograms were fitted with one exponential, and the closed-time distributions were fitted with the sum of two exponentials. The inset above the closed-time distribution represents the same data and fitted curve plotted over a time interval of 0.5–8.0 ms. The time constants (τ) were as follows: for PHA open time, 0.37 ms; closed time 0.6 ms (inset) and 9.9 ms; burst duration, 2.6 ms. For $\text{Ins}(1,4,5)\text{P}_3$ open time 0.34 ms; closed time 0.6 ms (inset) and 11.5 ms; burst duration 2.1 ms.



of the T cell by mitogen. Results suggested that stimulation apparently occurs through the mediation of an internal second messenger. The mitogen-operated channel has properties in common with the $\text{Ins}(1,4,5)\text{P}_3$ -activated conductance, including permeability to Ca^{2+} and Ba^{2+} ions, low single channel conductance and positive E_{rev} , kinetic behaviour characterized as brief openings in bursts, and inactivation by internal Ca^{2+} concentrations $\geq 10^{-6}$ M. To clarify whether $\text{Ins}(1,4,5)\text{P}_3$ may serve as the internal second messenger regulating transmembrane Ca^{2+} flux following mitogen stimulation, we compared the inward currents induced by the mitogen phytohaemagglutinin (PHA) in a cell-attached patch, to inward currents activated by $\text{Ins}(1,4,5)\text{P}_3$ in

the same patch after isolation to the inside-out recording configuration.

The results are shown in Fig. 3. Single-channel records from one patch, (1) cell-attached, before and after bath application of PHA, and (2) inside-out, before and after bath application of $\text{Ins}(1,4,5)\text{P}_3$ are shown in Fig. 3a. The cell was suspended in isotonic KCl, so transmembrane patch potential was assumed to be 0 mV in both the cell-attached and inside-out configurations. The similar appearance of the PHA-elicited and $\text{Ins}(1,4,5)\text{P}_3$ -activated conductances, characterized as intermittent bursts of brief openings with low amplitudes, suggested that they were the same channel. Comparative quantitative

analysis confirmed this impression. Probability density histograms of unitary current amplitude (Fig. 3b) were virtually identical. In each case, interval analysis revealed monoexponentially distributed open times and burst durations, and double exponentially distributed closed times; the kinetic parameters derived from open-time, closed-time and burst duration histograms were similar. Kinetic parameters are presented for comparative purposes only, as the limited signal-to-noise ratio precludes accurate quantitative kinetic analysis. These results strongly suggest that the $\text{Ins}(1,4,5)\text{P}_3$ -activated channel is the same as the one turned on during PHA-induced mitogenesis, and thus that $\text{Ins}(1,4,5)\text{P}_3$ is the internal second messenger mediating Ca^{2+} channel activation following mitogen-stimulation of T lymphocytes.

There is experimental precedence. Receptor-mediated activation of T lymphocytes is accompanied by prompt, sustained increases in phosphoinositide turnover³²⁻³⁷, in $\text{Ins}(1,4,5)\text{P}_3$ ³⁸ and in free cytosolic Ca^{2+} ions^{19,39-42}. Experimental evidence obtained with permeabilized Jurkat cells suggests that the initial, transient increase in Ca^{2+} is due to $\text{Ins}(1,4,5)\text{P}_3$ -mediated release of Ca^{2+} from intracellular stores³⁸. However, both $\text{Ins}(1,4,5)\text{P}_3$ and Ca^{2+} continue to rise during the first 10 min after T-cell activation and are still elevated at 30 min. It has previously been shown that the mitogen-regulated Ca^{2+} -permeable channel becomes active as early as 20 s after addition of mitogen and has an augmented opening probability for at least 10 min following T-cell activation. A reasonable hypothesis supported by the experimental data is that T-cell activation by mitogens elicits $\text{PtdIns}(4,5)\text{P}_2$ breakdown to the second messenger, $\text{Ins}(1,4,5)\text{P}_3$, which then acts both to release internal Ca^{2+} stores for the initial transient increase in Ca^{2+} and to activate the transmembrane Ca^{2+} -permeable channel for more sustained increases in Ca^{2+} . Evidence for $\text{Ins}(1,4,5)\text{P}_3$ -initiated transmembrane Ca^{2+} flux is also strong in blowfly salivary glands^{43,44} and clonal rat pituitary cells⁴⁵. Alternatively, preliminary experiments with rat liver plasma membrane vesicles indicate that inositol 1,3,4,5-tetrakisphosphate ($\text{Ins}(1,3,4,5)\text{P}_4$), which is derived by 3-phosphorylation of $\text{Ins}(1,4,5)\text{P}_3$ by a soluble kinase^{46,47}, promotes transmembrane Ca^{2+} fluxes. The generation of inositol polyphosphates which act as second messengers mediating transmembrane Ca^{2+} flux may be a mechanism common to a variety of other cell types exhibiting the phosphatidylinositol turnover response, whereby activation of cell-surface receptors brings about a sustained rise in cytosolic Ca^{2+} concentration.

This work was supported in part by American Cancer Society Institutional Research Grant IN-32Z and in part by a Grant-in-Aid from the American Heart Association with funds contributed in part by the Santa Clara County Chapter of the American Heart Association, California Affiliate and by the American Heart Association, Florida Affiliate. M.K. is supported by Dean's Postdoctoral Fellowship Funds, Stanford University. P.G. is the recipient of the Faculty Development Award of the Pharmaceutical Manufacturers' Association. We thank R. Dewart for secretarial assistance.

Received 15 August; accepted 31 December 1986.

- Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
- Hokin, L. E. *A. Rev. Biochem.* **54**, 205-235 (1985).
- Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67-69 (1983).
- Burgess, G. M. *et al.* *Nature* **309**, 63-66 (1984).
- Joseph, S. K., Thomas, A. P., Williams, R. J., Irvine, R. F. & Williamson, J. R. *J. biol. Chem.* **259**, 3077-3081 (1984).
- Burgess, G. M., Irvine, R. F., Berridge, M. J., McKinney, J. S. & Putney, J. W. *Biochem. J.* **232**, 237-243 (1985).
- Berridge, M. J., Heslop, J. P., Irvine, R. F. & Brown, K. D. *Biochem. J.* **222**, 195-201 (1984).
- Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **222**, 269-272 (1984).
- Biden, T. J., Prentki, M., Irvine, R. F., Berridge, M. J. & Wollheim, C. B. *Biochem. J.* **223**, 467-475 (1984).
- Prentki, M., Wollheim, C. B. & Lew, P. D. *J. biol. Chem.* **259**, 13777-13782 (1984).
- Gershengorn, M. C., Geras, E., Purrello, V. S. & Rebecchi, M. J. *J. biol. Chem.* **259**, 10675-10681 (1984).
- Hirata, M., Suematsu, E., Hashimoto, T., Hamachi, T. & Koga, T. *Biochem. J.* **223**, 229-236 (1984).
- Suematsu, E., Hirata, M., Hashimoto, T. & Kuriyama, H. *Biochem. biophys. Res. Commun.* **120**, 481-475 (1984).
- Dawson, A. P. & Irvine, R. F. *Biochem. biophys. Res. Commun.* **120**, 858-864 (1984).

- Streb, H., Bayerdorffer, E., Haase, W., Irvine, R. F. & Schulz, I. *J. Membrane Biol.* **81**, 241-253 (1984).
- Prentki, M., Janjic, D., Irvine, R. F., Berridge, M. J. & Wollheim, C. B. *Nature* **309**, 562-564 (1984).
- Michell, R. H. *Biochem. biophys. Acta* **415**, 81-147 (1975).
- Kuno, M., Goronzy, J., Weyand, C. M. & Gardner, P. *Nature* **323**, 269-273 (1986).
- Weiss, A., Imboden, J., Shoback, D. & Stobo, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4169-4173 (1984).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-200 (1981).
- Decoursey, T. E., Chandy, K. G., Gupta, S. & Cahalan, M. D. *Nature* **307**, 465-468 (1984).
- Matteson, D. R. & Deutsch, C. *Nature* **307**, 468-471 (1984).
- Cahalan, M. D., Chandy, K. G., Decoursey, T. E. & Gupta, S. *J. Physiol., Lond.* **358**, 197-237 (1985).
- Cavalié, A., Ochi, R., Pelzer, D. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **398**, 284-297 (1983).
- Hagiwara, S. & Ohmori, H. *J. Physiol., Lond.* **336**, 649-661 (1983).
- Chesnoy-Marchais, D. *J. Physiol., Lond.* **367**, 457-488 (1985).
- Matteson, D. R. & Armstrong, C. M. *J. gen. Physiol.* **87**, 161-182 (1986).
- Thomas, A. P., Alexander, J. & Williamson, J. R. *J. biol. Chem.* **259**, 5574-5584 (1984).
- Rittenhouse, S. E. & Sasson, J. P. *J. biol. Chem.* **260**, 8657-8660 (1985).
- Weiss, S. J., McKinney, J. S. & Putney, J. W. Jr. *Biochem. J.* **206**, 555-560 (1982).
- Rebecchi, M. J. & Gershengorn, M. C. *Biochem. J.* **216**, 287-294 (1983).
- Fisher, D. B. & Mueller, G. C. *Proc. natn. Acad. Sci. U.S.A.* **60**, 1396-1402 (1968).
- Fisher, D. B. & Mueller, G. C. *Biochem. biophys. Acta* **248**, 434-448 (1971).
- Allan, D. & Michell, R. H. *Biochem. J.* **164**, 389-397 (1977).
- Hui, D. Y. & Harmony, J. A. K. *Biochem. J.* **192**, 91-98 (1980).
- Hasegawa-Sasaki, H. & Sasaki, T. *J. Biochem.* **91**, 463-468 (1982).
- Hasegawa-Sasaki, H. & Sasaki, T. *Biochim. biophys. Acta* **754**, 305-314 (1983).
- Imboden, J. B. & Stobo, J. D. *J. exp. Med.* **161**, 446-456 (1985).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68-71 (1982).
- Weiss, M. J., Dazley, J. F., Hodgdon, J. C. & Reinherz, E. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6836-6840 (1984).
- Oetting, H. C., Terhorst, C., Cantley, L. C. & Rosoff, P. M. *Cell* **40**, 583-590 (1985).
- Nisbet-Brown, E., Cheung, R. K., Lee, J. W. W. & Gelfand, E. W. *Nature* **316**, 545-547 (1985).
- Berridge, M. J. & Fain, J. N. *Biochem. J.* **178**, 59-69 (1979).
- Fain, J. N. & Berridge, M. J. *Biochem. J.* **180**, 655-661 (1979).
- Albert, P. R. & Tashjian, A. H., Jr. *J. biol. Chem.* **259**, 15350-15363 (1984).
- Irvine, R. F., Letcher, A. J. & Heslop, J. P. *Nature* **320**, 631-634 (1986).
- Hansen, C. A., Mah, S. & Williamson, J. R. *J. biol. Chem.* **261**, 8100-8103 (1986).
- Marty, A. *Nature* **291**, 497-500 (1981).
- Sachs, F., Neil, J. & Barkakati, N. *Pflügers Arch. ges. Physiol.* **395**, 331-340 (1982).

HLA class II induction in human islet cells by interferon- γ plus tumour necrosis factor or lymphotoxin

Ricardo Pujol-Borrell*, Ian Todd*, Mala Doshi*, Gian Franco Bottazzo*, Robert Sutton†, Derek Gray‡, Gunther R. Adolf§ & Marc Feldmann§

* Immunology Department, Middlesex Hospital Medical School, London W1P 9FG, UK

† Nuffield Department of Surgery, John Radcliffe Hospital, Oxford OX3 9DU, UK

‡ Ernst-Boehringer Institute, A-1121, Vienna, Austria

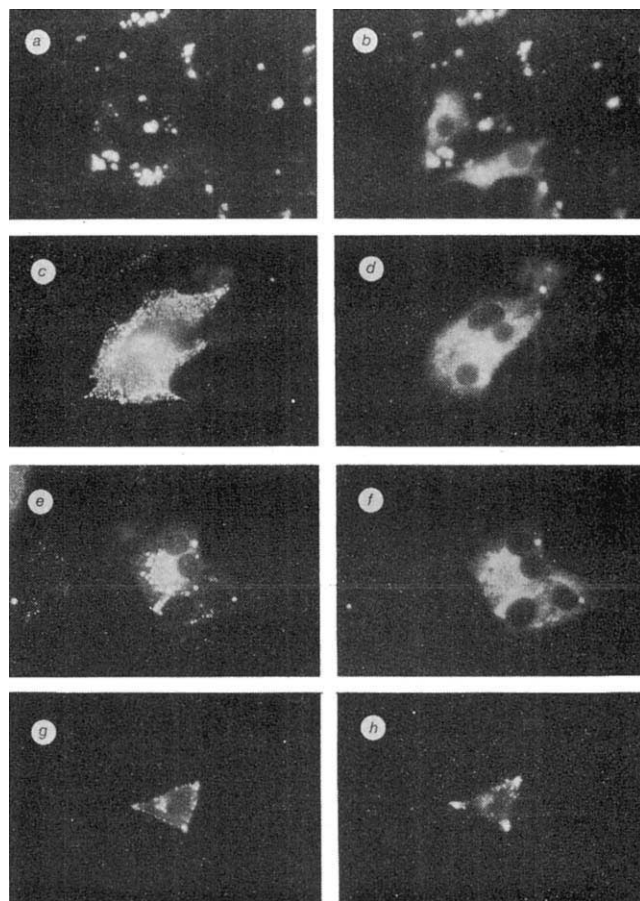
§ Charing Cross Sunley Medical Research Centre, Lurgan Avenue, London W6 8LW, UK

HLA class II molecules are surface glycoproteins which are essential in the initiation of immune responses. It has been postulated that induction of class II in epithelial cells such as endocrine cells, which are normally class II negative¹, may result in autoimmunity². In type I diabetes, islet beta cells, the target of the autoimmune process, selectively express class II antigens^{3,4}. But in contrast to most other cell types⁵, islet beta cells are not stimulated to express class II by interferon- γ (IFN- γ)^{6,7} and thus the conditions under which this induction occurs have been particularly elusive. The cytotoxins tumour necrosis factor (TNF) and lymphotoxin (LT) synergize with IFN- γ in a number of activities⁸. We report here that IFN- γ in combination with either TNF or LT induces islet cell class II expression. This finding has important implications for the pathogenesis of type I diabetes and the understanding of the differential control of class II expression.

The observations that normal thyroid follicular cells (thyrocytes) can be induced to express class II antigens⁹ and that these products are spontaneously present in thyrocytes from glands afflicted by autoimmune disease (Graves' thyrotoxicosis or Hashimoto's thyroiditis)¹⁰ led us to propose a new concept of the pathogenesis of cell-specific autoimmunity². According

Fig. 1 Indirect immunofluorescence micrographs ($\times 700$) of 8-day-old islet-cell-enriched pancreatic cultures stained three days after the addition of IFN- γ (50 U ml^{-1}) and TNF (100 U ml^{-1}). Cells were stained to detect membrane (a, c, g) and cytoplasmic (e) HLA class II antigens (fluorescein left-hand photography) and C-peptide (b, d, f) or glucagon (h) (rhodamine right-hand photography). Cultures which were not stimulated with IFN- γ and TNF or LT were negative when stained for class II (a) but contained cells expressing C-peptide (b). By contrast, in the stimulated cultures (c-h), a group of three C-peptide-positive beta cells (c) show positive granular surface membrane staining for class II (c); of a cluster of four beta cells positive for C-peptide (f) two show cytoplasmic class II positive staining (e); one isolated alpha cell showing class II membrane staining (g) was counterstained by mAb to glucagon (h).

Methods. Pancreatic cultures, from 8 different glands obtained at surgery from chronic pancreatitis ($n=2$) and from kidney donors ($n=6$) were prepared following the method recently described²⁸ modified when necessary to facilitate the processing of smaller amounts of tissue⁶. After digestion and filtration through meshes of different diameter, aliquots of 10^5 cells were plated on glass coverslips placed in 24-well plates (Nunc) with 0.5 ml of RPMI 1640 culture medium containing 15% fetal calf serum (FCS) (Sera Lab), 2 mM glutamine, 5 mg ml^{-1} transferrin (Sigma) and 10^{-8} M hydrocortisone (Sigma). At day 3 to 6 the medium was changed, FCS concentration lowered to 5% and the different supplements such as IFN- γ , TNF and LT added. IFN- γ (specific activity $\sim 3 \times 10^7$ U per mg protein; endotoxin contamination less than 0.125 ng per mg protein) and TNF (specific activity 5×10^7 U per mg protein; endotoxin contamination less than 0.125 ng per mg protein) were prepared from recombinant sources, whereas LT (specific activity $1-2 \times 10^7$ U per mg protein) was purified from supernatants of RPMI 1788 cells to a final purity of at least 90%²⁹. Double immunofluorescent staining of the cell monolayer was carried out 2-8 days after the addition of the supplements following a protocol previously described⁶. After extensive washing in Hanks balanced salt solution containing 1% BSA viable monolayers were first incubated for 30 min at room temperature with immunoglobulin (IgM) mouse mAb RFDR1 to the non-polymorphic region of the class II molecules (G. Janossy, Royal Free Hospital, London) washed again and subsequently stained with fluorescein isothiocyanate (FITC)-labelled affinity-purified goat anti-mouse (IgM) (μ -chain specific, Southern Biotechnology). Cultures were further washed and fixed for 5 min in 50% v/v acetone/methanol to expose cytoplasmic hormonal contents and incubated with the corresponding anti-pancreatic-hormone specific IgG mouse mAb (PEP 030: anti C-peptide; Som 018: anti-somatostatin; Gluc 1: anti-glucagon, all from Dr P. N. Jorgensen, Novo, Copenhagen). TRITC-(tetramethyl rhodamine isothiocyanate) labelled affinity-purified goat anti-mouse IgG (γ -chain specific, Southern Biotechnology) was then used to develop the reaction. Non-specific binding was ruled out by staining parallel cultures with an unrelated mAb of the same immunoglobulin class or normal mouse serum at dilution of 1:500. Undesirable cross-reactivities among reagents from different species were assessed by omitting one of the layers in turn. All preparations were examined with $\times 63$ oil-immersion objective under a fluorescence Zeiss III photomicroscope equipped with epi-illumination.



to this hypothesis, the transition from a local non-specific inflammatory reaction to autoimmunity would occur as a consequence of the aberrant expression of HLA class II molecules by the epithelial endocrine cells. In genetically susceptible individuals these cells could then efficiently present their own surface autoantigens to autoreactive helper T lymphocytes, bypassing the requirement for classical antigen-presenting cells. This concept is strongly supported by our observations that class II-expressing thyrocytes can present exogenous peptide¹¹ or endogenous autoantigens¹² to T-cell clones. The applicability of this hypothesis to type I diabetes has not yet been ascertained, but is suggested by the occurrence of inappropriate class II expression in the pancreatic islet beta cells of diabetic patients (refs 3 and 4; see ref. 13 for review).

Using the double immunofluorescence technique on monolayer cultures of human pancreas enriched in islet cells we previously showed that IFN- γ induces strong class II expression in pancreatic exocrine and ductal cells but not in the islet beta cells⁶. However, IFN- γ increases class I expression in human beta cells^{6,7} demonstrating the presence of IFN- γ receptors on these cells. When human pancreatic cells were similarly cultured with IFN- γ and TNF or IFN- γ and LT for 2-8 days, up to 50% of the beta cells showed positive membrane and cytoplasmic staining for class II, of intensity comparable to that seen in the ductal cells (Fig. 1). IFN- γ at 10 U ml^{-1} was sufficient when used together with 50 U ml^{-1} or more of either

of the cytotoxins. Conversely, 10 U ml^{-1} of either cytotoxin was effective in conjunction with doses of IFN- γ of at least 50 U ml^{-1} (Fig. 2).

The HLA class II expression induced *in vitro* on the beta cells was detected with various monoclonal antibodies (mAbs) including DA6.164, specific for the HLA-DR β -chain (Dr K. Cuy, Edinburgh) and VIC-Y1, specific for the cytoplasmic invariant chain (Professor W. Knapp, Vienna). A limited number of alpha and delta cells were identified in the cultures by using the anti-glucagon mAb Gluc-1 and the anti-somatostatin mAb SOM018, in the double immunofluorescence technique: both of these cell types also expressed class II when cultured with either of the combinations of IFN- γ with TNF or LT (Fig. 1g and h).

Interleukin-1 (IL-1) shares many of the biological effects of TNF¹⁴ and has been shown to influence beta-cell function¹⁵. We accordingly tested the effect of r(recombinant)IL-1 α and rIL-1 β (P. Lomedico, Hoffmann LaRoche and C. Lewis, Glaxo) alone or together with IFN- γ on class II expression by islet cells following the same protocol used for TNF. However both types of IL-1 were ineffective in the dose range $10-1,000 \text{ U ml}^{-1}$. The number of beta cells in the monolayers was unaffected by IL-1. The specificity of the effects of the TNF and LT preparations was established by observing $>80\%$ reduction in the number of class II-positive beta cells in cultures supplemented with 100 U ml^{-1} of IFN- γ together with TNF or LT pre-incubated with neutralizing antisera (see legend to Fig. 2c and d).

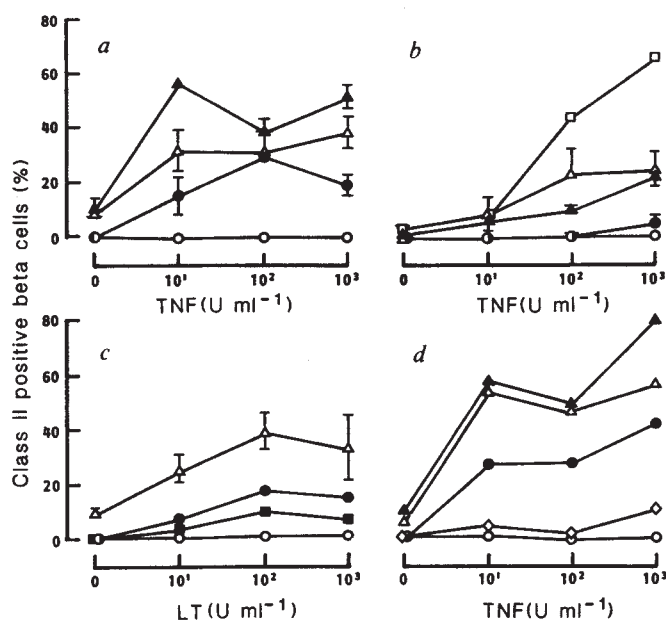


Fig. 2 Synergistic effect of IFN- γ and TNF (a, b, d) or LT (c) on beta cell class II expression. Pancreatic cell cultures were simultaneously supplemented at day 2 to 6 and tested 3–7 days later. a, b, c, Mean of data (bars $\pm$ s.e.m.) obtained from nine different sets of cultures in which cells from eight different glands were used. d, Results from an experiment carried out with cultures from a single pancreas. Graphs a, c, d refer to membrane, and b to cytoplasmic, class II expression. Symbols $\circ$, $\bullet$, Δ and $\blacktriangle$; 0, 10, 100 and 1,000 U ml $^{-1}$ of IFN- γ respectively. In b, $\square$ indicates cytoplasmic class II expression as detected by mAb VIC-Y1, directed against the invariant chain of class II molecules. In the absence of IFN- γ neither TNF nor LT induced class II expression in beta cells whereas IFN- γ by itself had a weak effect on a small percentage of beta cells (see text). In c, $\blacksquare$ indicates the reduced dose response to 1,000 U ml $^{-1}$ of IFN- γ and increasing doses of LT when these factors were preincubated for 4 h with rabbit antibodies to LT (batch 25C, final neutralizing capacity 1,000 U of LT ml $^{-1}$). In d, $\diamond$ indicates the similarly reduced dose response to 1,000 U ml $^{-1}$ of IFN- γ and increasing doses of TNF when these factors were preincubated with a rabbit antibody to TNF (batch 15-1, final neutralization capacity in the cultures 1,000 U TNF ml $^{-1}$).

In cultures supplemented with the cytotoxins alone, class II induction was not observed. IFN- γ alone at doses over 100 U ml $^{-1}$ occasionally produced weak class II expression in up to 10% of the beta cells, but the weakness of this staining led us to doubt its significance particularly when compared with that seen on the ductal cells in the same preparations⁶ or on human thyroid cultures⁷; others have reached similar conclusions⁷. The interpretation of these results is not clear but one possibility is that islets from certain donors were exposed *in vivo* to TNF or LT, thus rendering them weakly susceptible to induction by IFN- γ *in vitro*. Pancreatic monolayer cultures, though enriched in islet cells, contained over 80% non-endocrine cells among which ductal cells predominate but macrophages, dendritic-like cells and fibroblasts were also present. It is therefore possible that the effect of IFN- γ and TNF, or IFN- γ and LT, on the islet cells is exerted or potentiated through mediator(s) released by other cell types as shown in another system¹⁶. However, a direct action on the beta cells is favoured by the observation that the intensity of class II induction observed was similar in cultures in which the representation of the different cell types varied widely.

TNF and LT have recently been characterized and cloned^{17,18}. They show 30% sequence similarity, use the same receptor and synergize with IFN- γ in their cytotoxic action on transformed or malignant cells^{19,20}. Neither TNF nor LT, either alone or together with IFN- γ , seemed to have an acute cytotoxic effect

on beta cells as the number of these cells present in control and treated cultures did not differ significantly. The observations reported here constitute the first demonstration that TNF and LT have an influence on major histocompatibility complex (MHC) class II expression in human epithelial cells. These findings have wide implications for understanding how TNF and LT exert their immunomodulatory effects and stress the recently recognized similarities between the spectrum of action of these mediators and that of interleukins and interferons^{21–24}.

The finding of experimental conditions in which human beta cells can be induced to express HLA class II molecules not only strengthens the previous histopathological observations of the pancreas of diabetic individuals^{3,4} but also points to a possible sequence of events leading to this phenomenon *in vivo*. It has recently been shown that either TNF or LT in combination with IFN- γ can protect cells from viral infection although they kill those already infected⁸, so it is conceivable that the local release of these mediators in response to a viral infection of the beta cells is one of the mechanisms underlying inappropriate class II expression in the diabetic pancreas. The short half-life of these lymphokines should limit their sphere of action and thereby help to explain the beta cell specificity of the class II expression observed in the diabetic pancreas.

The results reported here demonstrate that multiple signals with a degree of cell and tissue specificity are necessary for the most efficient induction of class II expression. Such differential regulation is most dramatically illustrated with islet cells in which IFN- γ alone is virtually ineffective, but it also applies to other cell types^{25–27}. This synergy of mediators may limit the circumstances in which inappropriate induction of class II occurs and explain the marked heterogeneity of class II expression in different cell types. Furthermore, it indicates the possibility of therapeutic intervention in autoimmune diseases by selectively reducing class II expression in particular cell types, and avoiding the generalized immunosuppression that would result if class II expression were abrogated in all antigen-presenting cells.

We thank the scientists and firms cited for their gifts of reagents, our colleagues Mr Fernando, Mr Russell, Drs C. Brook, D. Katz, A. Kurtz and J. Rode (Middlesex Hospital) Dr Clark (St Bartholomew's Hospital, London), Dr D. Grant (Institute for Child Health, London) and Ms B. Watson for help in procuring pancreatic tissue, Professor Deborah Doniach, I. M. Roitt and P. J. Morris for support and encouragement and Marian Pine for editing and typing the manuscript. R.P.B. is supported by the Juvenile Diabetes Foundation International (USA) and I.T. by the Wellcome Trust.

Received 24 November 1986; accepted 29 January 1987.

1. Baekkeskov, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6456–6460 (1981).
2. Bottazzo, G. F., Pujol-Borrell, R., Hanafusa, T. & Feldmann, M. *Lancet* **ii**, 1115–1119 (1983).
3. Bottazzo, G. F. *et al. New Engl. J. Med.* **313**, 353–360 (1985).
4. Foulis, A. K. & Farquharson, M. A. *Diabetes* **35**, 1215–1224 (1986).
5. Todd, I. *et al. Clin. exp. Immun.* **61**, 265–273 (1985).
6. Pujol-Borrell, R. *et al. Clin. exp. Immun.* **65**, 128–139 (1986).
7. Campbell, I. L. *et al. J. clin. Endocr. Metab.* **62**, 1101–1109 (1986).
8. Wong, G. W. & Goeddel, D. V. *Nature* **323**, 819–822 (1986).
9. Pujol-Borrell, R., Hanafusa, T., Chiovato, L. & Bottazzo, G. F. *Nature* **303**, 71–73 (1983).
10. Hanafusa, T. *et al. Lancet* **ii**, 1111–1115 (1983).
11. Londei, M., Lamb, J. R., Bottazzo, G. F. & Feldmann, M. *Nature* **312**, 639–641 (1984).
12. Londei, M., Bottazzo, G. F. & Feldmann, M. *Science* **228**, 85–89 (1985).
13. Pujol-Borrell, R. & Todd, I. *Balliere's Clin. Immun. Allergy* (in the press).
14. Beutler, B. & Cerami, A. *Nature* **320**, 584–588 (1986).
15. Bendtzen, K. *et al. Science* **232**, 1545–1547 (1986).
16. Walker, E. B. *et al. J. exp. Med.* **159**, 1532–1547 (1984).
17. Gray, P. W. *et al. Nature* **312**, 721–724 (1984).
18. Pennica, D. *et al. Nature* **312**, 724–729 (1984).
19. Stone-Wolf, D. S. *et al. J. exp. Med.* **159**, 828–843 (1984).
20. Lee, S. H. *et al. J. Immun.* **133**, 1083–1086 (1984).
21. Shalaby, M. R. *et al. J. Immun.* **135**, 2069–2073 (1985).
22. Philip, R. & Epstein, L. B. *Nature* **323**, 86–89 (1986).
23. Pober, J. S. *J. Immun.* **137**, 1893–1896 (1986).
24. Collins, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 446–450 (1986).
25. Todd, I., McNally, J. M., Hammond, L. J. & Pujol-Borrell, R. in *Frontiers in Thyroidology* (eds Gaitan, E. & Mederos-Neto, G.) 1551–1554 (Plenum, New York, in the press).
26. Acres, R. B., Lamb, J. R. & Feldmann, M. *Immunology* **54**, 9–16 (1985).
27. Todd, I., Pujol-Borrell, R., Hammond, L. J., McNally, J. M., Feldmann, M. & Bottazzo, G. F. *Clin. exp. Immun.* (submitted).
28. Gray, D. W. R., McShane, P., Grant, B. & Morris, P. J. *Diabetes* **33**, 1055–1061 (1984).
29. Adolf, G. R. & Fogy, I. *Lymphokine Res.* **3**, 231–237 (1984).

Dominance of one T-cell receptor in the H-2K^b/TNP response

Ute Hochgeschwender*†, Hans-Georg Simon*†,
Hans Ulrich Weltzien†, Frauke Bartels*,
Andrea Becker* & Jörg T. Epplen*

* Junior Research Unit and † Cellular Immunology Unit,
Max-Planck-Institut für Immunbiologie, D-7800 Freiburg, FRG

T-cell receptors (TCRs) recognize foreign antigens in the context of major histocompatibility complex (MHC)-encoded cell surface proteins. These receptors are heterogeneous, dimeric glycoproteins composed of disulphide linked α - and β -chains. We analysed the diversity of TCRs in a collection of H-2K^b-restricted, 2,4,6-trinitrophenyl (TNP)-specific (H-2K^b/TNP) cytotoxic T-cell (T_c) clones from C57BL/6 mice. Investigation of the β -chain messenger RNAs revealed that nearly half of these independent clones expressed an identical β -chain gene¹. We show here that almost all the T_c clones expressing the predominant β -chain gene also express an identical α -chain gene. These results show that a strong selective pressure acted on the T_c population, resulting in a skewing of the TCR repertoire for H-2K^b/TNP and in the dominant expression of one TCR with this specificity. Possible explanations for this skewing include antigen-driven clonal expansion and network interactions.

Vertebrate T cells are capable of recognizing a multitude of foreign antigens in the context of MHC-encoded cell-surface proteins². The diversity of TCR specificities results from the somatic rearrangement of germline gene segments for both α - and β -chains³. The same mechanisms for generating immunoglobulin diversity (such as germline diversity, junctional diversity, combinatorial joining and combinatorial association) also generate diversity for TCRs, with the possible exception of somatic mutation⁴. The α - and β -chains of TCR alone confer antigen and MHC specificity on T cells⁵. The antigen and MHC specificities are properties of the TCR as a whole rather than separate properties of the α - and β -chains^{5,6}. As yet, no general correlations between TCR structure and function have been observed.

Although the potential number of TCRs is high, the actual number expressed in T-cell populations with a given specificity is unknown⁷. The analogous problem of immunoglobulin repertoire diversity has been studied extensively⁸, but the differences between the functions of immunoglobulins and TCRs preclude extension of these results to TCRs. The restriction of the TCR repertoire may reflect specific aspects of T-cell function. For instance, TCRs are selected for restriction to recognizing antigen in the context of one self-MHC encoded protein and for tolerance to all self-MHC encoded proteins⁹⁻¹¹. Further, possible preferential expansion of some T cells expressing a given specificity or interactions of idiotypic networks might also alter TCR repertoire.

To address the question of the diversity of the TCR repertoire, we have analysed the TCR genes in a panel of 42 independent

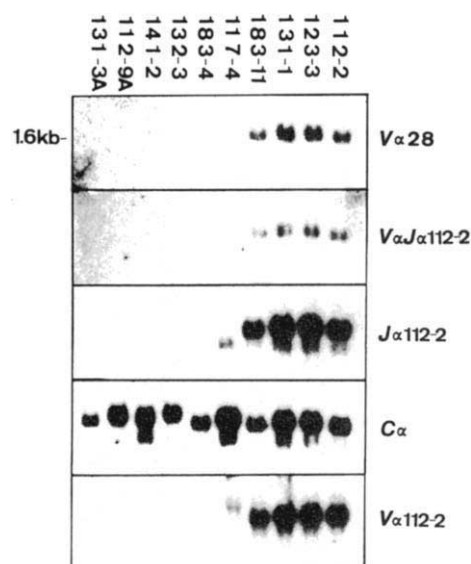


Fig. 2 Analysis of T-cell receptor α -chain mRNAs with oligonucleotide probes. RNA from 10 different T_c clones specific for H-2K^b/TNP were hybridized with five different ³²P-labelled oligonucleotides specific for different TCR gene segments. For description of oligonucleotide probes see Fig. 1 and for details of methods see refs 1, 22 and 23.

T_c clones specific for H-2K^b/TNP¹². Previously we reported that almost half of these clones express identical β -chain genes¹. The high frequency of this β -chain might result from its ability to combine with a large number of different α -chains to generate H-2K^b/TNP specificity. Alternatively, the β -chain may be associated with only one or a few related α -chains, indicating a strong selection for this TCR idiomorph. Paradigms for both possibilities exist in the studies of B-cell repertoire^{13,14}. To answer this question we have now analysed the α -chain mRNAs of these T_c clones.

We cloned a partial cDNA copy of the α -chain mRNA from one of the T_c clones expressing the predominant β -chain mRNA, T_c clone 112-2. The cDNA was made from total RNA of T_c clone 112-2 and inserted into the bacteriophage vector λ gt11. This library was screened with an oligonucleotide probe specific for the constant region of the α -chain gene. Strongly hybridizing clones were isolated and the cDNA inserts sequenced. One clone had an insert of 650 bp which consisted of the 3' portion of V α , a J α and the 5' portion of C α (Fig. 1). On the basis of this sequence we synthesized oligonucleotide probes (see Fig. 1). The V α 112-2 probe is complementary to the V α gene segment expressed in T_c clone 112-2 and two other sequences in the genome (data not shown). The V α 28 probe is complementary to the 3' end of the V α 112-2 and of the V α 28 gene segments (nomenclature according to ref. 15). The probe specific for J α 112-2 is a previously unreported J α gene segment expressed in T_c clone 112-2. Finally, the V α J α 112-2 probe is complementary to the junction between the V α 112-2 and J α 112-2 gene segments.

Fig. 1 Nucleotide sequence of the V α J α junction in the H-2K^b/TNP-specific T_c clone 112-2 and derived synthetic oligonucleotide probes.

Methods. A cDNA library was constructed from total cellular RNA of T_c clone 112-2 in the bacteriophage vector λ gt11¹⁹ as described¹. Screening was carried out using an oligonucleotide probe (5'

cDNA	5' ...TTGAAGCCACATACAATAAAGAAACACCTCCTCCACTGTCAGAAAGCCTCAGTGCAAGAGTCAGACTCGGCTGTGT...	V
Oligo-probes	3' TCGGTGTATGTTATTTCTTTGGT 5' 3' CGTCTTTCCGAGTCACGTTCTCA 5'	V α 112-2 V α 28
cDNA	...ACTACTGTGCTCTGGGTGATCCCTATGCCAGGGATTAACCTTCGGTCTTGGCACCAGAGTATCTGTGTTCCCTACA...	V J C α
Oligo-probes	3' TGACACGAGACCACTAGGGATA 5' 3' AAGCCAGAACCGTGGTCTCATA 5'	V α J α 112-2 J α 112-2

GTGCTGTCTGAGACCGAGGATC 3') specific for the C α region²⁰. Three independent positive cDNA phage clones with inserts of 500–650 base pairs were subcloned into pUC12 and M13mp9. Sequence analysis was carried out according to Sanger *et al.*²¹ with universal and specific primers. The sequences consist of more than 100 base pairs of the 3' end of a V α , J α and a nearly full-length C α element.

α -Chain gene element				Designation of T _C clones	β -Chain gene element			
V α 112-2	V α 28	V α J α 112-2	J α 112-2		V β 3	D β 112-2	J β 2-6	C β 2
				112-2				
				112-3				
				117-3				
				123-1				
				123-3				
				131-1				
				147-3				
				162-1				
				163-2				
				177-1				
				182-5				
				183-2				
				183-6				
				183-9				
				183-11				
				187-2				
*		*		122-1				
*		*		122-6				
*		*		133-2				
				112-6A				
				117-1				
				117-4				
				131-2				
				132-4				
				127-1				
				183-3				
				183-3.1				
				183-4				
				183-7				
				132-3				
				141-2				
				161-1				
				163-4				
				183-1				
				112-9A				
				113-4A				
				117-2A				
				121-1				
				131-3A				
				142-2A				
				143-2A				
				147-1				

Fig. 3 T-cell receptor α - and β -chain genes of H-2K^b/TNP-specific T_C clones. T_C clones are listed according to their gene structure as determined by Northern blot or RNA or DNA gel hybridizations with the oligonucleotide probes. For detailed description of the clones see refs 1 and 12. From three T_C clones (122-1, 122-6, 133-2) neither RNA nor DNA was available to complete the results for hybridization with the V α 112-2 and the V α J α 112-2 probes. *, Not tested.

These oligonucleotides were used to probe the RNAs of the panel of T_C clones, either by Northern blotting or by direct gel hybridizations. In Fig. 2 we show a representative example of the hybridization results; the results of all the RNA hybridizations are summarized in Fig. 3. With one exception (T_C clone 133-2) all the T_C clones expressing the predominant β -chain mRNA express α -chain mRNAs composed of identical germline segments. All the rearrangements of V α 112-2 to J α 112-2 gener-

ated very similar, probably identical, V α J α junctions as evidenced by the hybridization of the V α J α 112-2 probe. Eight of the sixteen T_C clones expressing this $\alpha\beta$ dimer were shown to be independent on the basis of having different γ -chain or non-productive β -chain rearrangements (T_C clone 112-2, 117-3, 123-3, 131-1, 147-3, 177-1, 183-2, 183-6; data not shown). Thus these T_C clones are not sister isolates but rather independent examples of identical α - and β -chain gene rearrangements. Of the T_C clones not expressing the predominant β -chain mRNA, one T_C clone (117-1) expresses an α -chain mRNA composed of V α 112-2 fused to another J α and T_C clone 131-2 has a different V α fused to J α 112-2. Finally, two T_C clones (117-2A, 147-1) express mRNA which hybridizes with the V α 112-2 probe but not with the V α 28 probe and thus express a related V α element.

The oligonucleotide probes were also used in hybridizations with B6 liver and T_C clone DNAs. As expected, the oligonucleotide probe specific for the junction V α J α 112-2 does not hybridize to the germline DNA but only to DNA of those T_C clones expressing the V α 112-2/J α 112-2 mRNA. The migration of the V α J α 112-2 hybridizing band is coincident with the migration of a V α 112-2 and a J α 112-2 hybridizing non-germline band in these T_C clones (data not shown). Although both V α 112-2 and V α 28 probes are complementary to multiple fragments in liver DNA, in all 16 T_C clones expressing the same V α gene the identical V α element is rearranged.

As the sequence of V α 112-2 has not been published, we isolated the B6 liver genomic fragment containing V α 112-2. The sequence is shown in Fig. 4: V α 112-2 has similarity with the V α 4 family¹⁵ (versus V α 28: 79% in the nucleic acid sequence and 77% in the predicted protein sequence; versus V α 65: 78% in the nucleic acid sequence and 73% in the predicted protein sequence).

In summary, 40% (16/42) of independent T_C clones isolated for H-2K^b/TNP specificity express α - and β -chain genes composed of identical germline segments and with identical joints. This result has two significant aspects: that one TCR can dominate the immune response and that the sequence of this TCR is highly conserved in independent isolates. The high degree of sequence conservation is surprising in the light of the work of Fink *et al.*⁴. In analysing five T helper-cell clones specific for H-2IE^k/pigeon cytochrome c, they found that the TCRs were related, but varied significantly in D β , J β and J α sequences. These differences correlated with altered antigen cross-reactivity and alloreactivity. The findings suggest that in this system significant variation in α - and β -chain genes can still result in the same antigen/MHC specificity. Our own finding that the predominant β -chain may complex with either the predominant α -chain or with a related α -chain (T_C clone 133-2) to confer the same specificity support this view. Thus we suspect that the conservation of the predominant α - and β -chain gene sequences in our system is not required for antigen/MHC specificity, but reflects its selection at one or several stages in T-cell development.

For T cells expressing one TCR to dominate the response to a given antigen/MHC combination, they must have had a selec-

```

CTGCAGGCTCACAATAGATGCCAGGCATATAAAGGTGTGTTATTTCTATGTATGCTTAGAGCTGTAATCTGAAGAAACAACTTGAGATGCTTTCC
100
GTTTAAAGTCAGACAGTGTCTCTATGTTGAAGCTTATTGTAAGCAGGATAGTTCCAAAATCAGAGCTATACCTTGATAGCTGTCTTTGGACA
200
TAAAGGACTGAAAGACTGACATACACTTTATTGGTAAGATCAGGAGAGAGCCCACTGAAATATCCATCCTCAGTTACGCTCCTGCCACAGGGGGCGC
300
AGATGCTTAACACTGGCTTTGTTCTGGGTGGTTTATGCTTAAGATGACTCAGAGTTTCTCCGAGTTAACACTCCTTCATGCTAAGCATCAAGACCACT
400
M T L K M D S S P G F V A V I L L I L
TCTAGATGACACTAAAGATGGACTCTTCTCCAGGCTCTGGGTGTGATCTCTCATACTTGGTAAGTAAACAGAGTATTCTTTAGCGGTAAAGTTATCT
500
TCAAGGCAACAGAGCTTATCTCCCTCTCTCTTCTCCCTTCCCTTTGCACTGTAGGAAGGACCCAGGAGATTCGGTGACTCAACAGAGGCCAGT
600
T V S E S E S L I I N C T Y S A T S I A Y P N L F W Y V R Y P G E
GACCGTCTCAGAAAGCGAGTCCCTGATAATAATGTCACGTATTACGCCACAGCATAGCTTACCCCTAATCTTTCTGGTATGTCGATATCTGGAGAA
700
G L Q L L L K V I T A G O K G S S R G F E A T Y N K E T T S F H L
GGTCTACAACCTCTCTGAAAGTCATTACGGCTGGCCAGAGGGAAGCAGCAGAGGGTTGAAGCCACATACATAAAGAAACCACTCTCTCCACTTGC
800
O K A S V Q E S D S A V Y Y C A L G
AGAAAGCCTCAGTCAAGAGTCAGACTCGGTGTGTACTACTGTCTCTGGTGACACAGTGGCAGAACTGCAG

```

Fig. 4 Nucleotide sequence of the unrearranged V α 112-2 gene segment. The predicted amino-acid sequence encoded by the leader and the V region exons is identified by the single letter code. A library of B6 DNA in Charon 28²⁴ was screened with V α 112-2 and V α 28 oligonucleotide probes. Positive clones were isolated, subcloned into M13mp9, and sequenced according to Sanger *et al.*²¹.

tive growth advantage over other T cells with the same specificity. This advantage might be expressed at a number of stages in T-cell development. The rearrangement of TCR germline segments may not be random. An example of non-random rearrangement exists in the choice of immunoglobulin V gene segments early in B-cell ontogeny¹⁶. The selection for self-MHC restriction and tolerance may favour T cells expressing certain receptors⁹⁻¹¹. During the immune response itself the stimulation with antigen may preferentially expand the population of T cells expressing certain TCRs. The basis for this preferential stimulation need not be higher affinity as is seen in the maturation of B-cell responses⁸. Finally, the TCR may confer an advantage in the context of an idiotype network¹⁷. Indeed, T cells expressing anti-TCR idiotype specificity have been isolated¹⁸.

Our study shows that a significant skewing of the TCR repertoire occurs in the response to H-2K^b/TNP. Further work will be necessary to determine the generality of dominant TCRs in T-cell responses and the exact nature of the selection for these TCRs.

We thank Uli Birsner for synthesizing and purifying the oligonucleotides. The computer comparisons were kindly performed by R. B. Wallace using the program COMPARE (DNASar). The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00263.

Received 27 October 1986; accepted 20 January 1987.

- Hochgeschwender, U., Weltzien, H. U., Eichmann, K., Wallace, R. B. & Epplen, J. T. *Nature* **322**, 376-378 (1986).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-177 (1979).
- Kronenberg, M., Siu, G., Hood, L. E. & Shastri, N. A. *Rev. Immun.* **4**, 529-591 (1986).
- Fink, P. J., Matis, L. A., McElligott, D. L., Bookman, M. & Hedrick, S. M. *Nature* **321**, 219-226 (1986).
- Dembić, Z. *et al. Nature* **320**, 232-238 (1986).
- Kappler, J. W., Skidmore, B., White, J. & Marrack, P. J. *exp. Med.* **153**, 1198-1214 (1981).
- Goverman, J., Hunkapiller, T. & Hood, L. *Cell* **45**, 475-484 (1986).
- Rudikoff, S. *Cont. Top. molec. Immun.* **6**, 169-209 (1983).
- Zinkernagel, R. M. *Immun. Rev.* **42**, 224-270 (1978).
- Heber-Katz, E. *et al. J. exp. Med.* **155**, 1086-1099 (1982).
- Hünig, T. R. & Bevan, M. J. *J. exp. Med.* **155**, 111-125 (1982).
- Weltzien, H. U., Kempkes, B., Jankovic, D. L. & Eichmann, K. *Eur. J. Immun.* **16**, 631-639 (1986).
- Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. E. *Nature* **283**, 35-40 (1980).
- Kaartinen, M., Griffiths, G. M., Markham, A. F. & Milstein, C. *Nature* **304**, 320-324 (1983).
- Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
- Yancopoulos, G. D. & Alt, F. W. *A. Rev. Immun.* **4**, 339-368 (1986).
- Jerne, N. K. *Ann. Immun.* **125**, 373-389 (1974).
- Lamb, J. R. & Feldman, M. *Nature* **300**, 456-458 (1982).
- Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194-1198 (1983).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Sanger, F., Coulson, A. R., Barrel, B. G., Smith, A. J. H. & Roe, B. J. *J. molec. Biol.* **143**, 161-178 (1980).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Tsao, S. G. S., Brunk, C. F. & Pearlman, R. E. *Analyt. Biochem.* **131**, 365-372 (1983).
- Blattner, F. R. *et al. Science* **196**, 161-169 (1977).

Possible role for fatty acyl-coenzyme A in intracellular protein transport

Benjamin S. Glick & James E. Rothman

Department of Biochemistry, Stanford University, Stanford, California 94305, USA

The transport of proteins between subcellular compartments is a vectorial, energy-requiring process mediated by the budding and fusion of a series of vesicular carriers¹. As yet, nothing is known of the chemical reactions that underlie these events, or how or in exactly what forms energy is used to sustain such movements. Here we report that fatty acyl-CoA acts as cofactor to a Golgi-associated protein factor (termed NSF) that is required for transport between cisternae of the Golgi stack in a cell-free system. This previously unsuspected connection may offer a link between the complex process of protein transport and a single, well-defined type of chemical reaction. We suggest that an ATP-dependent cycle of fatty acylation and deacylation may play an important role in driving rounds of vectorial protein transport.

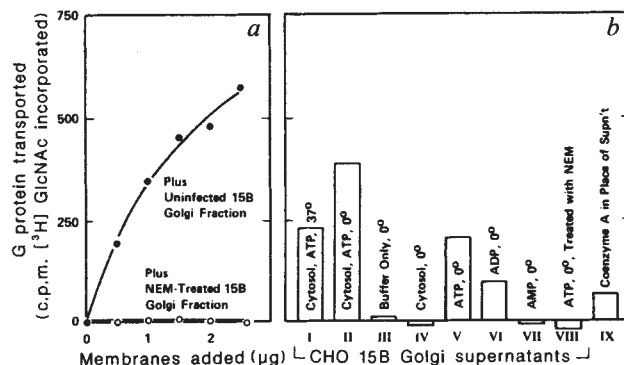


Fig. 1 Restoration of transport activity of NEM-treated donor and acceptor by CHO 15B Golgi fraction. *a*, G protein transport with (○) NEM-treated or (●) untreated uninfected CHO 15B Golgi fraction added to assay mixture, in which donor and acceptor had both been treated with NEM, before incubation at 37 °C. *b*, G protein transport in the same assay mixture, containing supernatants of CHO 15B Golgi fraction that had been incubated with various combinations of CHO 15B cytosol and nucleotides at 37 °C (I) or 0 °C (II-VIII). In mixture IX, CoASH replaced CHO 15B Golgi fraction supernatant.

Methods. G protein transport assays were performed as described previously³. Reaction mixes (50 μl total volume) contained donor and acceptor Golgi membranes, gel-filtered CHO 15B cytosol, 50 μM ATP (standard concentration), an ATP-regenerating system, and assay buffer consisting of 25 mM HEPES-KOH (pH 7.0), 15 mM KCl, and 2.5 mM magnesium acetate. Incubations were for 60 min at 37 °C. Cytosol was prespun 30 min at 100,000g in a Beckman Airfuge to eliminate any membrane contaminants. NEM was freshly dissolved to a concentration of 50 mM, and diluted 1:50 into Golgi fractions containing 1.0 M sucrose, 10 mM Tris (pH 7.4). After 15 min at 0 °C, an equal volume of 100 mM dithiothreitol (DTT) was added to quench unreacted NEM. Immunoprecipitation of G protein was carried out overnight at 4 °C, as DTT inhibited immune complex formation at 37 °C. Protein was measured according to Lowry¹⁹ using bovine serum albumin (BSA) as a standard. In *a*, background incorporation of 143 c.p.m. (obtained in the absence of uninfected 15B Golgi) has been subtracted. For comparison, 787 c.p.m. above background were incorporated into G protein when NEM-treated donor was assayed together with untreated acceptor. Uninfected CHO 15B Golgi were pretreated with NEM (1 mM) by incubating for 15 min at 0 °C followed by 2 mM DTT before addition. In *b*, uninfected CHO 15B membranes (0.42 mg ml⁻¹) were diluted fivefold into assay buffer. Where indicated, 1.6 mg ml⁻¹ CHO 15B cytosol and/or 100 μM nucleotide were also present. Sample I was incubated 15 min at 37 °C; all others were kept on ice. Samples (100 μl each) were spun for 10 min at 400,000g in a Beckman TL-100 tabletop ultracentrifuge. Samples V and VIII were identical except that the supernatant from sample VIII was treated with NEM (1 mM, 15 min at 0 °C, followed by 2 mM DTT). Portions (25 μl) of the supernatant were assayed for restorative activity. Sample IX contained CoASH (10 μM final concentration) in place of an uninfected 15B membrane supernatant. A background of 223 c.p.m. (with no uninfected 15B supernatant present) has been subtracted. Sample II (25 μl) before centrifugation (containing 2.1 μg CHO 15B membranes) gave 1,076 c.p.m. above background.

The combined requirements for NSF and fatty acyl-coenzyme A were revealed by their ability to restore transport in a cell-free system whose membrane components had been selectively inactivated with the sulphhydryl reagent *N*-ethylmaleimide (NEM). This cell-free system reconstitutes the transport of the vesicular stomatitis virus (VSV)-encoded glycoprotein (G protein) between successive cisternae of the Golgi stack²⁻⁷. This transport reaction requires ATP as well as a cytosol (high-speed supernatant) fraction. G protein starts out in a Golgi-enriched 'donor' fraction from a VSV-infected mutant CHO cell line (clone 15B) that lacks the enzyme *N*-acetylglucosaminyl-transferase I. This Golgi fraction is incubated with a comparable Golgi fraction (termed 'acceptor') prepared from uninfected

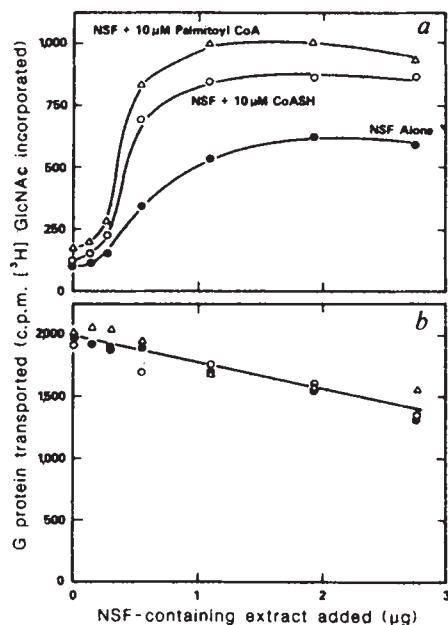


Fig. 2 Stimulation by CoASH and palmitoyl-CoA of the restorative activity of NSF under standard assay conditions. Donor and acceptor membranes were either (a) treated with NEM (1 mM, 15 min at 0 °C, followed by 2 mM DTT), or (b) not treated. ●, NSF alone; ○, NSF plus 10 μM CoASH; △, NSF plus 10 μM palmitoyl-CoA. Transport assays were performed as in Fig. 1, except the amounts of the components added were reduced by half to give a reaction volume of 25 μl. The NSF-containing fraction was prepared by adding 4.75 ml of uninfected CHO 15B membranes (0.82 mg ml⁻¹) to 19 ml of ice-cold assay buffer containing 100 μM ATP (final concentration). Membranes were pelleted 60 min at 300,000g in a Beckman SW-50.1 rotor. The supernatant was concentrated 14-fold using Amicon Centricon 30 microconcentrators, and stored frozen at -80 °C. The active concentrate contained 0.55 mg ml⁻¹ protein (measured according to Bradford²⁰, using bovine serum albumin (BSA) as a standard). CoASH and palmitoyl-CoA (Sigma) were added from aqueous stock solutions.

wild-type CHO cells. When G protein is transported to the *N*-acetylglucosamine-transferring (GlcNAc transferase-containing) medial⁸ cisternae of the wild-type acceptor Golgi stacks, in an incubation containing UDP-[³H]GlcNAc, then [³H]GlcNAc is incorporated into G protein, providing a measure of transport during the cell-free incubation. Current evidence suggests that this transport is mediated by a population of non-clathrin coated vesicles⁶.

Pretreatment of both donor and acceptor membrane fractions with NEM (1 mM for 15 min at 0 °C) completely inhibited the transport reaction in assays containing untreated cytosol⁹. However, transport was virtually normal if only the donor or the acceptor fraction was pretreated. This suggested that an NSF-sensitive factor (NSF) was needed for transport, and that NSF could be supplied by either the donor or the acceptor Golgi membranes, each of which would be expected to contain this factor. Because the requirement for NSF was evident in the presence of untreated cytosol, NSF activity cannot be present in appreciable amounts in the cytosol fraction as prepared.

To test this interpretation, we examined whether the Golgi fraction from uninfected CHO 15B cells could restore transport activity to reactions containing NEM-treated donor and NEM-treated acceptor membranes (Fig. 1a). Golgi fractions prepared from uninfected CHO 15B cells possess neither donor activity (as they contain no G protein) nor acceptor activity (as they lack the GlcNAc transferase), and thus do not normally have a stimulatory effect on the assay. However, uninfected CHO 15B Golgi membranes completely restored transport activity to

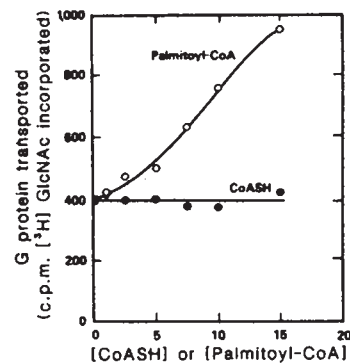


Fig. 3 Effect of palmitoyl-CoA (○) and CoASH (●) on transport at low ATP levels. Final concentration of ATP was 0.5 μM; other conditions were as described in Fig. 1. As the CoASH and acyl-CoA effects are strongest when NSF is present but in limiting amounts, endogenous NSF on the membranes was partially inactivated. (This cannot easily be achieved by adding NSF back to completely NEM-inactivated membranes, because ATP is present in the soluble NSF preparations described.) Partial inactivation was obtained with a relatively mild treatment of donor and acceptor (0.5 mM NEM for 1 min at 0 °C) after which G protein transport is reduced by about 80%, versus the 95% inhibition typically seen using 1 mM NEM for 15 min.

mixtures of NEM-treated donor and NEM-treated acceptor (Fig. 1a). This restorative activity is our definition of NSF, which is conveniently assayed because the entire transport process is dependent upon this factor. Presumably, the NSF added in the form of Golgi membranes replaces the endogenous NSF that had been quenched by NEM treatment of donor and acceptor. As expected, pretreatment of the uninfected CHO 15B membranes with NEM abolished their ability to restore transport (Fig. 1a).

The ability of uninfected CHO 15B membranes to provide NSF suggested that this factor was being released from membranes under the conditions of our transport reaction. To find conditions to extract NSF from the Golgi membranes in a soluble form, high-speed supernatants obtained after various incubations of the CHO 15B Golgi membranes were assayed for their restorative activity (Fig. 1b). About 20% of the restorative activity of uninfected CHO 15B Golgi membranes was released into the supernatant after incubation with ATP at 0 °C. Neither the cytosol fraction nor incubation at 37 °C was required for NSF dissociation (Fig. 1b). Release of NSF occurred within 5 min at 0 °C and depended on the presence of a nucleotide such as ATP (Fig. 1b). Half-maximal release occurred at about 65 μM ATP. ADP, but not AMP, could replace ATP in promoting release (Fig. 1b). (Other nucleotides have not been tested.) The significance of the rapid, temperature-independent release of NSF by a nucleotide is not yet clear, but this behaviour suggests that NSF is peripherally bound to the cytoplasmic surfaces of membranes in the Golgi fraction.

This ATP-extracted supernatant was used as the standard preparation of NSF. The soluble form of NSF retained its NEM-sensitivity (1 mM, 0 °C, 15 min) (Fig. 1b). The relative molecular mass of NSF is greater than 100,000 as determined by its ultrafiltration characteristics, and it is inactivated by trypsin and by boiling (Table 1). These results imply that NSF is a macromolecule, partially or entirely proteinaceous.

While testing a series of well-known small molecules containing sulphhydryl groups as controls to illustrate the specificity of the restorative activity of NSF (which contains an essential sulphhydryl), we were struck by the fact that reduced coenzyme A had some restorative activity, although much less than could be obtained with the bona fide (proteinaceous) NSF (Fig. 1b). Moreover, CoASH (at 10 μM) greatly stimulated the ability of NSF to restore transport activity to NEM-treated membranes

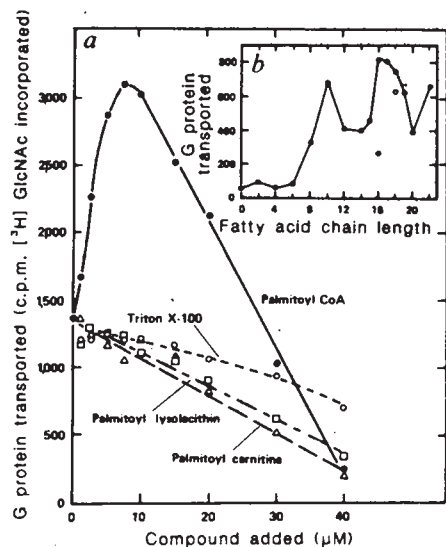


Fig. 4 Comparison of stimulation by acyl-CoA with the effects of detergents. Donor and acceptor were partially inactivated with NEM (0.5 mM for 1 min at 0 °C) as in Fig. 3. *a*, Transport assays were performed as in Fig. 1 (with 50 μM ATP). Palmitoyl-CoA (●), Triton X-100 (○), palmitoyl lysocleithin (Δ) and palmitoyl carnitine (□) were added as aqueous solutions to the concentrations indicated. *b*, The carbon chain length dependence of stimulatory activity was tested for various saturated fatty acyl-CoAs (●) and the monounsaturated palmitoleoyl (C_{16:1}) CoA and oleoyl (C_{18:1}) CoAs (○). Chain length 0 corresponds to free CoASH. Fatty acyl-CoAs were added as aqueous solutions to a final concentration of 15 μM. ATP concentration was 0.5 μM. A background of 477 c.p.m. (obtained with no acyl-CoA present) has been subtracted. All reagents were from Sigma.

(Fig. 2*a*). CoASH was not simply acting as a nonspecific reductant, as the CoASH effect was manifest in the presence of excess (0.4 mM) dithiothreitol (Fig. 2*a*) and also because reduced glutathione was totally inactive in replacing CoASH (not shown). Note that the synergy observed with CoASH is greatest with low concentrations of NSF (Fig. 2*a*). Also, neither CoASH nor NSF significantly stimulated control transport reactions containing untreated donor and acceptor membranes, in which saturating amounts of endogenous NSF are present (Fig. 2*b*). These data suggest that CoASH (or some derivative of CoASH) and NSF act together in a step required for protein transport.

The major function of CoA is as a carrier of activated acyl groups, linked by a thioester bond. Desulpho CoA (in which the SH is replaced by H) was inactive (data not shown), suggesting that the CoA effect involves the linkage of CoA through a thioester bond to endogenous acyl groups. If so, then an acyl-CoA, rather than CoASH itself, would be the true cofactor to NSF. As shown in Fig. 2*a*, palmitoyl-CoA replaced CoASH, and offered an even better stimulation of NSF activity. Palmitoyl-CoA did not appreciably stimulate the complete transport reaction in which NSF is not limiting (Fig. 2*b*).

That palmitoyl-CoA was more active than CoASH on a molar basis suggested that this fatty acyl-CoA was acting directly, rather than indirectly by supplying CoASH upon breakdown. The apparent cofactor activity of CoASH would then be the result of the action of endogenous fatty acid-CoA ligase, which catalyses the reaction: endogenous fatty acid + CoASH + ATP $\rightleftharpoons$ fatty acyl-CoA + AMP + PP_i. In the absence of this ligase activity, fatty acyl-CoA should still serve as cofactor to NSF, but CoASH should no longer have any cofactor activity. To test this prediction, we took advantage of the fact that fatty acid-CoA ligase (from rat liver microsomes) has a *K_m* for ATP of about 2–5 mM (refs 10, 11), whereas the transport reaction has an apparent *K_m* for ATP of only about 1 μM (ref. 9) (although we

Table 1 Inactivation of NSF by high temperature and proteolysis

NEM-treated donor and NEM-treated acceptor plus:	[³ H] GlnNAc incorporated into G protein (c.p.m.)
A, No addition	283
B, NSF, preincubated 10 min at 37 °C before assay	3,474
C, NSF, boiled 3 min at 100 °C	237
D, NSF, trypsin-treated, then trypsin inhibitor	238
E, NSF, trypsin inhibitor, then trypsin	2,288

Transport assays were performed as in Fig. 1. Concentrated NSF (1.0 mg ml⁻¹ protein) was prepared as described in Fig. 2. For samples B and C, 2 μl of water were added to 10 μl of NSF. For sample D, 1 μl of 1 mg ml⁻¹ trypsin (Sigma, type XIII) was added to 10 μl of NSF. After 5 min at 37 °C, 1 μl of 2 mg ml⁻¹ soybean trypsin inhibitor (Sigma) was added, followed by a further 5 min at 37 °C. For sample E, the order of addition of trypsin and trypsin inhibitor was reversed. Sample B was placed at 37 °C for 10 min to control for the effects of incubation. Five μl of samples B–E were assayed for NSF activity. Sample E was less active than sample B presumably because the inhibition of trypsin was not instantaneous.

Table 2 Dependence of transport of the unacylated VSV New Jersey G protein on NSF and acyl-CoA

Donor + Acceptor	[³ H] GlnNAc incorporated into G protein (c.p.m.)
NEM-Donor + NEM-Acceptor plus:	1,473
A, No addition	184
B, NSF	379
C, Palmitoyl-CoA	237
D, NSF + Palmitoyl-CoA	687
E, NEM-treated NSF + Palmitoyl-CoA	204

Donor membranes were prepared exactly as described³, except that the infection was done with the New Jersey strain of VSV. Transport assays were performed as in Fig. 1. Immunoprecipitation was carried out using a rabbit antiserum raised against a synthetic peptide corresponding to the carboxy-terminal 15 amino acids of the New Jersey G protein (Thomas Chappell, unpublished). Staph A cells (0.4 ml; Pansorbin, Calbiochem Corp.) were precoated with antibody (0.2 ml serum), and 10 μl of a 20% suspension were added to 50 μl incubations in place of the anti-G protein antiserum we normally use for immunoprecipitation. Where indicated, 1.3 μg concentrated NSF (prepared as described in Fig. 2) and/or 10 μM palmitoyl-CoA were added. For sample E, NSF was treated with 1 mM NEM for 15 min at 0 °C, followed by 2 mM dithiothreitol (DTT).

routinely use 50 μM ATP). Therefore, at very low ATP concentrations, ligase activity should be selectively reduced relative to transport. We found that palmitoyl-CoA retained its cofactor activity but CoASH no longer stimulated NSF in assays containing 0.5 μM ATP (Fig. 3). Presumably, transport at such low ATP levels is normally driven by an endogenous pool of acyl-CoA molecules. These data strongly suggest that fatty acyl-CoA itself acts as cofactor to NSF in promoting transport, although it is formally possible that some other kind of endogenously activated acyl group is enzymatically exchanged for the acyl moiety of acyl-CoA to form the active species. Acyl-CoAs with a broad range of fatty-acid chain lengths had cofactor activity when measured at 0.5 μM ATP to prevent resynthesis (Fig. 4*b*). The activity of short chain derivatives (such as C₈–C₁₂) may be significant, but might also be due to protection of endogenous pools of long-chain acyl-CoA from enzymatic hydrolysis.

What might be the mechanism of action of fatty acyl-CoA in promoting protein transport in the cell-free system? One possibility is that acyl-CoA acts as a bulk detergent¹². This seems unlikely because although palmitoyl-CoA stimulates NSF-

dependent transport, palmitoyl carnitine, palmitoyl lysolecithin and Triton X-100 do not stimulate, but rather inhibit transport, even at the lowest concentrations tested (Fig. 4a). Palmitoyl-CoA eventually inhibits the transport reaction (Fig. 4a), presumably because of a bulk detergent effect at high concentrations. The stimulation by long-chain acyl-CoAs is strongest at concentrations similar to those found in intact cells¹³. A second possibility is that acyl-CoA acts as an intact molecule in some specific fashion, perhaps as an allosteric effector of an enzyme¹⁴ or as a detergent when locally concentrated by a specific process.

We think the most attractive hypothesis, in line with its general role in cells, is that acyl-CoA acts by donating its fatty-acid chain to a recipient molecule, possibly a protein. Acylation could activate a component of the transport machinery; deacylation would then recycle this component for another round. NSF would be involved in this acylation-deacylation cycle. Energy in the form of ATP would be needed, at least through its use in fatty acyl-CoA synthesis. As G protein is known to be fatty acylated¹⁵, it is conceivable that G protein itself is the acyl acceptor in this postulated cycle. However, the fatty acylation of G protein has been shown to be unnecessary for transport *in vivo*¹⁶. Moreover, G protein is already acylated before it is transported in the cell-free system, and retains its fatty-acid chain during its transport¹⁷. To rule out directly any involvement of G protein acylation in the stimulation of G protein transport by NSF and acyl-CoA, we prepared donor membranes from CHO 15B cells that had been infected with the New Jersey strain of VSV. Unlike the Indiana strain we normally use, the New Jersey virus contains G protein that is not acylated *in vivo*¹⁸. As expected, treatment of Golgi membranes with NEM abolished

transport of New Jersey G protein, and the restorative activity of NSF was enhanced by palmitoyl-CoA (Table 2). Thus the putative fatty-acid recipient is not the transported protein itself.

Finding any small molecule that is needed for a complex cellular process can only help to pinpoint a well-defined, underlying chemical reaction. Although the exact meaning of the requirement for fatty acyl-CoA in protein transport is not yet clear, this result is likely to facilitate a biochemical dissection.

We thank Dr Felix Wieland for his advice, and Lyne Paquet for technical assistance. This work was supported by NIH grant GM25662-08, and an NSF Predoctoral Fellowship to B.S.G.

Received 21 July; accepted 5 November 1986.

1. Palade, G. E. *Science* **189**, 347-357 (1975).
2. Fries, E. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3870-3874 (1980).
3. Balch, W. E., Dunphy, W. G., Braell, W. A. & Rothman, J. E. *Cell* **39**, 405-416 (1984).
4. Braell, W. A., Balch, W. E., Dobbertin, D. C. & Rothman, J. E. *Cell* **39**, 511-524 (1984).
5. Balch, W. E., Glick, B. S. & Rothman, J. E. *Cell* **39**, 525-536 (1984).
6. Orci, L., Glick, B. S. & Rothman, J. E. *Cell* **46**, 171-184 (1986).
7. Wattenberg, B. W., Balch, W. E. & Rothman, J. E. *J. biol. Chem.* **261**, 2202-2207 (1986).
8. Dunphy, W. G., Brands, R. & Rothman, J. E. *Cell* **40**, 463-472 (1985).
9. Balch, W. E. & Rothman, J. E. *Arch. biochem. biophys.* **240**, 413-425 (1985).
10. Bar-Tana, J., Rose, G. & Shapiro, B. *Meth. Enzym.* **35**, 117-122 (1975).
11. Tanaka, T., Hosaka, K., Hoshimaru, M. & Numa, S. *Eur. J. Biochem.* **98**, 165-172 (1979).
12. Constantinides, P. P. & Steim, J. M. *J. biol. Chem.* **260**, 7573-7580 (1985).
13. Woldegiorgis, G. *et al. Anal. Biochem.* **150**, 8-12 (1985).
14. Edgar, J. R. & Bell, R. M. *J. biol. Chem.* **254**, 1016-1021 (1979).
15. Schmidt, M. F. G. *Curr. Top. Microbiol. Immun.* **102**, 101-124 (1983).
16. Rose, J. E., Adams, G. A. & Gallione, C. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2050-2054 (1984).
17. Dunphy, W. G., Fries, E., Urbani, L. J. & Rothman, J. E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7453-7457 (1981).
18. Gallione, C. J. & Rose, J. K. *J. Virol.* **46**, 162-169 (1983).
19. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
20. Bradford, M. M. *Anal. Biochem.* **72**, 248-254 (1976).

Stimulation of protein-directed strand exchange by a DNA helicase

Thomas Kodadek & Bruce M. Alberts

Department of Biochemistry and Biophysics,
University of California, San Francisco, San Francisco,
California 94143, USA

The protein-mediated exchange of strands between a DNA double helix and a homologous DNA single strand involves both synapsis and branch migration, which are two important aspects of any general recombination reaction¹. Purified DNA-dependent ATPases from *Escherichia coli* (recA protein)^{2,3}, *Ustilago* (rec 1 protein)⁴ and phage T4 (UvsX protein)⁵⁻⁷ have been shown to drive both synapsis and branch migration *in vitro*. The T4 gene 32 protein is a helix-destabilizing protein that greatly stimulates uvsX-protein-catalysed synapsis⁸, and the *E. coli* SSB (single-strand binding) protein stimulates the analogous recA-protein-mediated reaction to a lesser degree⁹. One suspects that several other proteins also play a role in the strand exchange process. For example, a DNA helicase could in principle accelerate branch migration rates by helping to melt the helix at the branch point. The T4 dda protein is a DNA helicase^{9,10} that is required to move the T4 replication fork past DNA template-bound proteins *in vitro*¹¹. Previously, we have shown that the dda protein binds to a column that contains immobilized T4 uvsX protein¹². We show here that this helicase specifically stimulates the branch migration reaction that the uvsX protein catalyses as a central part of the genetic recombination process in a T4 bacteriophage-infected cell.

The T4 uvsX protein catalyses branch migration in a directed manner, moving the branch point in the 5' to 3' direction with respect to the invading strand⁵. To monitor this reaction, we have employed the assay schematically illustrated in Fig. 1. A 3' end-labelled double-stranded linear M13 DNA molecule is incubated with a homologous single-stranded DNA circle in the

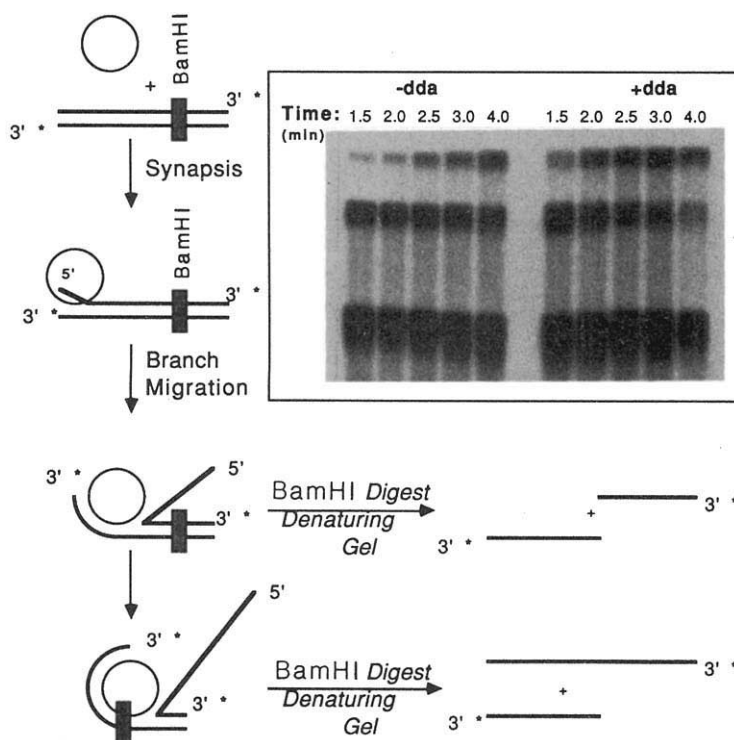
presence of the T4 uvsX and gene 32 proteins and ATP. A productive synapsis rapidly ensues at one end of the DNA molecule followed by a slower process of branch migration. To measure the extent of branch migration, aliquots are removed at various times and cut by brief exposure to a large excess of *Bam*HI restriction nuclease; the DNA products are then analysed by electrophoresis through a denaturing agarose gel^{13,14}. As *Bam*HI cuts the M13 DNA only once, two radioactive DNA bands that are less than full length are expected on autoradiography of the gel (one from each strand), so long as the heteroduplex region has not extended beyond the *Bam*HI site. However, as soon as the branch point moves past this site, the displaced radioactive strand will no longer be cut, and a full-length strand 6,407 nucleotides (nt) long will appear on the gel, replacing one of the two shorter DNA fragments (Fig. 1).

In the experiment shown in the insert in Fig. 1, the *Bam*HI site was located 2,220 base pairs (bp) from the end of the linear DNA molecule at which synapsis occurs, and about three minutes were required for the branch point to reach this site. As expected, shorter and longer times are required in this assay, respectively, if the *Bam*HI site is instead located 474 or 3,547 base pairs from this end (data not shown). We can thereby calculate a rate of branch migration of approximately 15 bp per second for the uvsX protein-catalysed process. Figure 1 also shows the effect of the T4 dda protein on this reaction. The addition of this protein was delayed until 1 min, a time that is sufficient to allow complete synapsis (data not shown). Within 0.5 min after its addition (1.5 min point), a considerable amount of the full-length DNA strand is observed. Quantitation reveals that the helicase has increased the rate of branch migration to approximately 70 bp s⁻¹, a stimulation of more than 4-fold.

In Fig. 2, autoradiograms of this kind have been scanned with a densitometer, and the degree of restriction nuclease resistance plotted as a function of time for experiments in which different amounts of the dda protein were added. As we have previously reported for replication fork movement¹⁵, the acceleration of branch migration increases with helicase concentration and all the DNA molecules in the population seem to display equivalent

Fig. 1 The rate of *uvrX* protein-catalysed branch migration is increased by the addition of the T4 *dda* protein, a DNA helicase. The design of the experiment is illustrated schematically at left and below; inset, autoradiogram showing results. In this experiment, the *Bam*HI site was located 2,220 nucleotides from the end at which synthesis is initiated; therefore, after DNA denaturation, radioactive DNA fragments that are 2,220 and 4,187 nucleotides long should be produced by the *Bam*HI restriction nuclease before the branch point reaches this site, whereas radioactive DNA fragments that are full-length (6,407 nucleotides) and 2,220 nucleotides long should be produced after the branch point passes the site. Although synthesis can begin at either end of the linear double helix, pairing at the opposite end to that shown is unproductive; the directional branch migration that follows will quickly separate the two interacting DNA molecules and prevent strand exchange.

Methods. Each reaction contained 2.5 μM (100 $\mu\text{g ml}^{-1}$) *uvrX* protein, 1.5 μM (50 $\mu\text{g ml}^{-1}$) gene-32 protein, 15.4 μM (as nucleotides) M13 single-stranded DNA, and 6.5 μM 3'- ^{32}P -labelled M13 linear double-stranded DNA (produced by treating the replicative form I (RFI) with *Hpa*I, followed by end-labelling with T4 DNA polymerase²² to a specific activity of 150 c.p.m. pmol^{-1}). The salts present were 10 mM Tris acetate (pH 7.4), 10 mM magnesium acetate, 1 mM dithiothreitol, 100 $\mu\text{g ml}^{-1}$ bovine serum albumin (BSA) and 90 mM potassium acetate. An ATP regeneration system (10 mM creatine phosphate and 1 U ml^{-1} creatine phosphokinase) was also present. Each reaction was preincubated at 37 °C for 2 min, and synthesis was then initiated by the addition of 2 mM ATP (time $t=0$). In the second experiment (+*dda*), the *dda* protein was added to a concentration of 0.07 μM (4 $\mu\text{g ml}^{-1}$) 1 min after the ATP. Aliquots of 10 μl were removed at the times indicated, and mixed with 24 units of *Bam*HI in 10 μl of restriction buffer (10 mM Tris-HCl, pH 8.1, 100 $\mu\text{g ml}^{-1}$ bovine BSA serum albumin, 300 mM NaCl, 150 μM poly(dT)); the elevated salt in this buffer dissociates the *uvrX* protein from double-stranded DNA). After 30 s at 37 °C, the reaction was quenched by adding NaOH to 30 mM and EDTA to 20 mM. The samples were then electrophoresed through a 1.2% alkaline agarose gel (30 mM NaOH) at 2 V cm^{-1} for 30 h. The gel was neutralized, dried onto DE-81 paper, and the radioactive DNA visualized by autoradiography. The top, middle and lower bands are 6,407, 4,187 and 2,220 nucleotides long, respectively.



rates. For DNA synthesis, this type of result has been attributed to the distributive mode of action of this DNA helicase, which seems to come on and off the DNA rapidly during the course of an experiment.

Branch migration is likely to be driven by an ATP dependent, 5' to 3' polymerization of *uvrX* protein monomers along the displaced single strand at the branch point (refs 14 and 16; T.K. and B.M.A., manuscript in preparation). The *dda* protein functions by first binding to a single strand overhang, and then moving along the DNA in a 5' to 3' direction. It therefore seems reasonable to propose that the *dda* protein also binds to the displaced strand near the branch point, and then moves rapidly into the helix (Fig. 3). The *dda* protein could thereby stimulate the growth of the *uvrX* protein filament, as illustrated. Because of its distributive mode of action, the *dda* protein presumably falls off the DNA after traversing less than a few hundred base pairs, and the cycle begins again.

Can the *dda* protein catalyse branch migration by itself? This possibility was tested in an experiment in which an aliquot of a reaction, initiated under our typical conditions, was diluted 10-fold into a solution identical to the original reaction mixture, except that the *uvrX* protein was omitted. At the resulting 10-fold lower concentration of the *uvrX* protein, the polar branch migration reaction that it catalyses is abolished (T.K. and B.M.A., in preparation). Adding 4 $\mu\text{g ml}^{-1}$ of the *dda* protein to the diluted mixture does not affect this result (data not shown). We conclude that the *dda*-protein-stimulated branch migration illustrated in Fig. 3 requires cooperation between the *uvrX* and *dda* proteins, and does not reflect a reaction of the *dda* protein alone.

Protein affinity chromatography studies have shown that the *dda* protein binds directly to both the *uvrX* and gene 32 proteins¹². Are either of these protein-protein contacts involved in

the observed stimulation of branch migration? To explore this question, we first asked whether the *dda* protein can stimulate the branch migration reaction that is catalysed by the *E. coli* *recA* protein. The *E. coli* *recA* protein is very similar in its activities to the T4 *uvrX* protein, and it is known to drive strand exchange with the same polarity as the *uvrX* protein³. Figure 4

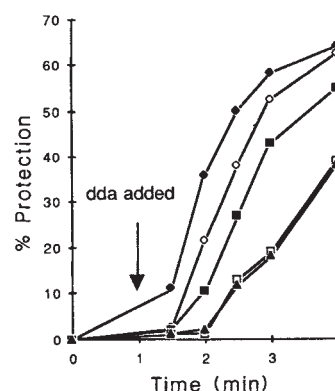


Fig. 2 The rate of branch migration increases with increasing concentration of *dda* protein. Reactions were conducted exactly as described in Fig. 1, except that the amount of the *dda* protein added was varied: $\blacklozenge$, 3.8 $\mu\text{g ml}^{-1}$; $\diamond$, 1.9 $\mu\text{g ml}^{-1}$; $\blacksquare$, 0.5 $\mu\text{g ml}^{-1}$; $\square$, none. $\blacktriangle$, Experiment in which the T4 gene 41 protein was added to a final concentration of 0.8 μM (50 $\mu\text{g ml}^{-1}$) in place of the *dda* protein and 1 mM GTP was provided. The autoradiograms of the denaturing agarose gels were scanned with a soft laser densitometer, and the percentage of protected molecules was plotted as a function of time.

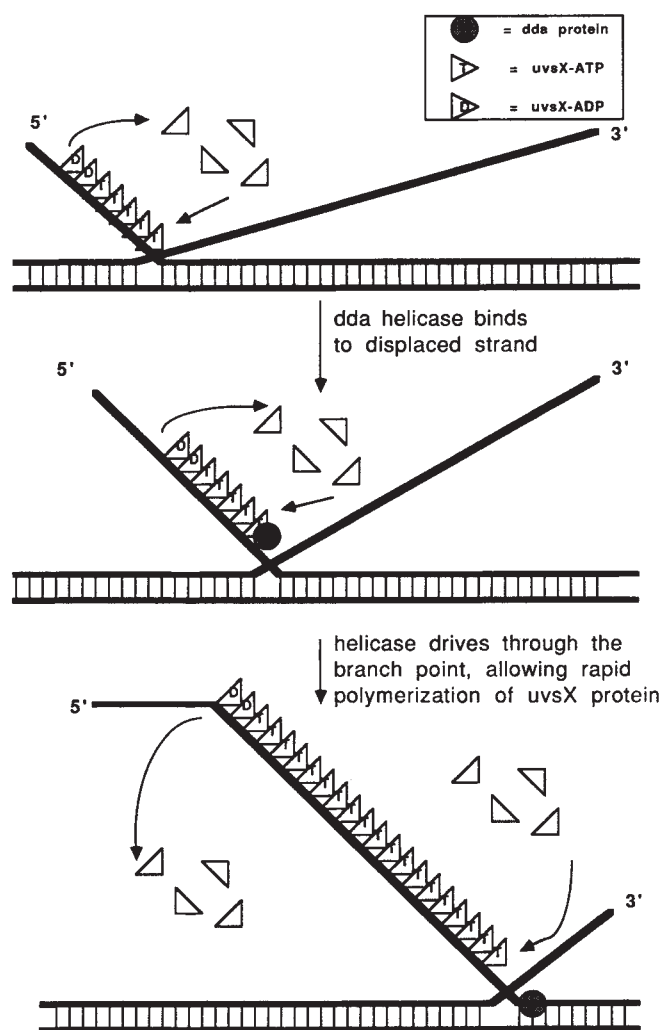


Fig. 3 A model for the stimulation of uvsX protein-driven branch migration by the dda protein (stippled circle). The helicase is loaded onto the displaced strand near the branch point, and moves rapidly into the double helix by translocating itself in a 5' to 3' direction along this strand using the energy of ATP hydrolysis (uvsX-ADP and uvsX-ATP are shown as arrowheads labelled D and T respectively); it thereby facilitates the directional assembly of an uvsX-protein filament and accelerates strand exchange. Efficient action of the dda helicase on the displaced strand may require a direct interaction of the dda protein with the uvsX protein, because the dda protein does not catalyse branch migration efficiently when the uvsX protein concentration is reduced (see text).

shows the result of an experiment similar to the one shown in Fig. 1, except that the *E. coli* recA and SSB proteins were substituted for their T4 analogues; presumably the dda protein does not bind to either of these *E. coli* proteins. In this experiment, the unique *Bam*HI site was located 3,547 bp from the end of the molecule, and significant protection from *Bam*HI nuclease was observed about 20 min after initiation of the reaction. This yields a branch migration rate of 3 bp s⁻¹ for the recA-protein-catalysed reaction, in good agreement with the value reported previously by Radding and coworkers¹⁴. In contrast to the uvsX protein-catalysed reaction, the dda helicase does not accelerate the rate of this reaction (Fig. 4).

Previous work has demonstrated that the T4 gene 41 protein, like the dda gene product, is a DNA helicase that moves along a DNA strand with 5' to 3' polarity^{17,18}. However, affinity chromatography suggests that the 41 protein does not directly interact with either the uvsX or gene 32 proteins^{12,19}. As shown in Fig. 2, no acceleration of the uvsX-protein-catalysed branch

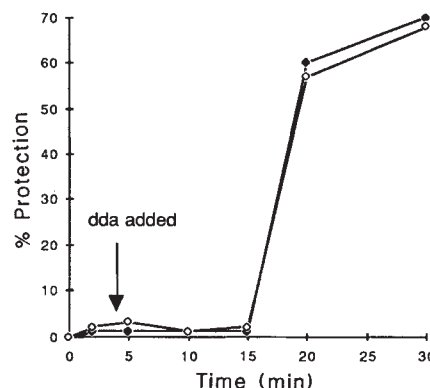


Fig. 4 The dda helicase has no effect on recA-protein-promoted branch migration. The reaction outlined in Fig. 1 was conducted according to Radding and coworkers¹⁴, and the *Bam*HI protection assay was employed to monitor the rate. In this experiment, the double-stranded DNA had been produced by *Bal*I treatment of the RFI M13 DNA, placing the *Bam*HI site 3,547 base pairs from the end at which synapsis begins. In the '+dda' reaction (◆), the dda protein was added at 0.07 μ M (4 μ g ml⁻¹) 4 min after addition of the ATP (which initiated pairing); ◇, no dda protein added. The autoradiogram of the gel was scanned, and the results plotted as in Fig. 2.

migration rate is detected when the 41 protein is added to the reaction. The same result is obtained when a high concentration of the T4 gene 61 product is added, which stimulates the helicase activity of the 41 protein¹⁸ and completes the T4 primosome²⁰ (data not shown).

The failure of the dda protein to stimulate the recA-protein-catalysed reaction, and of the gene 41 protein to accelerate the uvsX-protein-catalysed reaction, suggests that specific protein-protein interactions are involved in the observed stimulation of branch migration by the dda protein. As one possibility, this direct interaction could be required for efficient loading of the DNA helicase onto the displaced strand at the head of a uvsX nucleoprotein filament (Fig. 3).

It is becoming increasingly apparent that many complex processes in DNA metabolism, including transcription and replication, are mediated by highly specialized protein complexes that can be viewed as 'protein machines'²¹. The results reported here suggest that general recombination may involve catalytic machinery of comparable elegance.

We thank Dr Harold Selick for technical advice and helpful discussions. This work was supported by a grant from the National Institutes of Health and postdoctoral fellowship support from the Jane Coffin Childs Cancer Research Foundation (T.K.).

Received 18 December 1986; accepted 3 February 1987.

1. Radding, C. M. *A. Rev. Genet.* **16**, 405-437 (1982).
2. Flory, S. S. *et al. Cold Spring Harb. Symp. quant. Biol.* **49**, 513-523 (1984).
3. Cox, M. M. & Lehman, I. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3433-3437 (1981).
4. Kmiec, E. B. & Holloman, W. K. *Cell* **36**, 593-598 (1984).
5. Yonesaki, T. & Minagawa, T. *EMBO J.* **4**, 3321-3327 (1985).
6. Formosa, T. & Alberts, B. M. *J. biol. Chem.* **261**, 6107-6118 (1986).
7. Hinton, D. M. & Nossal, N. G. *J. biol. Chem.* **261**, 5663-5673 (1986).
8. Cox, M. M. *et al. J. biol. Chem.* **258**, 2577-2585 (1983).
9. Jongeneel, C. V., Formosa, T. & Alberts, B. M. *J. biol. Chem.* **259**, 12925-12932 (1984).
10. Krell, H., Dorwald, H. & Hoffmann-Berling, H. *Eur. J. Biochem.* **93**, 387-395 (1979).
11. Bedinger, P. *et al. Cell* **34**, 115-123 (1983).
12. Formosa, T. & Alberts, B. M. *Cold Spring Harb. Symp. quant. Biol.* **49**, 363-370 (1984).
13. Hsieh, P., Meyn, M. S. & Camerini-Otero, R. D. *Cell* **44**, 885-894 (1986).
14. Chow, S. A. *et al. J. biol. Chem.* **261**, 6961-6971 (1986).
15. Jongeneel, C. V., Bedinger, P. & Alberts, B. M. *J. biol. Chem.* **259**, 12933-12938 (1984).
16. Register, J. C. & Griffith, J. *J. biol. Chem.* **260**, 12308-12312 (1985).
17. Liu, C.-C. & Alberts, B. M. *J. biol. Chem.* **256**, 2813-2820 (1981).
18. Venkatesan, M., Silver, L. L. & Nossal, N. G. *J. biol. Chem.* **257**, 12426-12434 (1982).
19. Formosa, T., Burke, R. L. & Alberts, B. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2442-2446 (1983).
20. Alberts, B. M. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 655-668 (1983).
21. Alberts, B. M. *Cold Spring Harb. Symp. quant. Biol.* **49**, 1-12 (1984).
22. O'Farrell, P. *Focus* **3**, 1-3 (1981).

Amino-terminal processing of proteins

A. Ben-Bassat and K. Bauer

Recombinant proteins produced in bacteria often retain the initiation methionine at their amino-terminal end. This N-terminal methionine may affect the immunogenicity of the protein, necessitating removal.

PROTEIN synthesis is always initiated with either methionine in eukaryotes or N-formylmethionine in prokaryotes and eukaryotic organelles. The N-formyl moiety present in the proteins of eukaryotic organelles and bacteria is normally removed by deformylase, leaving methionine at the N-terminus. In the production of many intracellular proteins the methionine is removed by methionine aminopeptidase (MAP) after the initiation of translation. The substrate specificity and universality of MAP among prokaryotes and eukaryotes has been inferred from the N-terminal sequence analysis of mature cytosolic proteins^{1,2}. Recently the substrate specificity of MAP was verified using enzyme from *Escherichia coli*.³

Influential neighbours

Methionine cleavage from the N-terminal end of the protein is strongly affected by the adjacent amino-acid residue. When the side chain of this penultimate amino acid is large, hydrophobic or positively charged, the methionine is retained; when the side chain is small and uncharged the methionine is cleaved (Table 1). Consequently, the DNA sequence of a gene can be used to predict the fate of the N-terminal methionine of the resulting protein. These conclusions are based in part on qualitative sequence analysis of endogenous cytosolic proteins^{1,2}; partial removal of the methionine may be counted as a non-cleavage event. Partial processing of methionine can happen, especially when the penultimate amino acid is Ile, Val, Cys, or another amino acid of intermediate size.

Although no similar studies have been conducted with recombinant proteins, it is reasonable to assume that similar rules apply to their processing. Our experience with recombinant protein expression generally supports the conclusions obtained from endogenous cytosolic proteins. In addition, we found that other factors besides the penultimate residue affect the cleavage of methionine and that the cleavage of the methionine from hyperproduced recombinant proteins is frequently incomplete³.

Many of the recombinant proteins that are produced for therapeutic uses are based on secretory products from mammalian cells. The N-terminal signal peptides in these secretory precursors are cleaved during the translocation process; thus the mature proteins do not normally contain methionine at the N-terminal end. Examples are γ -interferon (Gln-Asp-),

tumour necrosis factor (Val-Arg-Ser-), and human growth hormone (Phe-Pro-). Production of these proteins in *E. coli* intracellularly often results in a product with an extra methionine. Obviously, the biotechnology industry has a strong interest in removing this methionine in order to avoid potential immunogenicity problems, and to obtain a more homogeneous product, without incompletely processed molecules.

Several methods have been used to remove the extra methionine. This first is removal of the methionine by treatment with cyanogen bromide⁴. This method involves chemical modification and requires extra preparation steps. It is not N-

Table 1 Cleavage of N-terminal methionine in intracellular prokaryotic proteins.

Met cleaved	Met removal varied	Met retained
Ala	Cys	Lys
Gly	His	Arg
Pro	Met	Leu
Ser	Trp	Ile
Thr	Tyr	Asn
	Asp	Phe
	Glu	
	Gln	
	Val	

The data show the effect of the penultimate amino acid on methionine removal².

terminal specific, and therefore its application is limited to proteins lacking internal methionines.

Fusion of the coding sequence of the recombinant protein with a sequence coding for a signal secretory peptide is another method⁵. In this procedure, the fused protein is secreted through the cytoplasmic membrane and cleaved at the fusion site by a signal peptidase. The recombinant protein can then accumulate in the periplasmic space in Gram-negative bacteria or in the growth medium in yeast and Gram-positive bacteria. Unfortunately, secretion of mammalian cell products in prokaryotes and yeast is often limited.

Enzymatic removal of the N-terminal methionine with aminopeptidase has also been used⁶. Here the methionine is removed by incubating the recombinant protein with a peptidase from *Aeromonas proteolytica*. This enzyme is a nonspecific aminopeptidase that cleaves amino acids sequentially from the N-terminal end of proteins. The stop signals for the enzyme are Glu, Asp and x-Pro, where x can be any amino acid except proline^{6,7}. This is suitable for removing N-terminal methionine when a stop signal is close to the methionine. The initiation methionine in

γ -interferon (Met-Glu-Pro-) and bovine growth hormone (Met-Asp-Glu-) have been removed using this enzyme. When the stop signals are not near methionine, proteolysis by this enzyme can continue, digesting the rest of the protein until it encounters a stop signal. Examples are human growth hormone analogue (Met-Lys-Val-Glu-), where methionine, lysine and valine are removed⁸.

The latest advance

The most recent development in N-terminal processing is the *in vivo* and *in vitro* removal of the N-terminal methionine using the MAP enzyme³. The MAP gene was recently isolated, cloned and overexpressed in *E. coli* by the authors and their coworkers. The *in vivo* method involves expression of the recombinant protein in a MAP-hyperproducing *E. coli* host. The advantage of this method is that it does not require additional purification steps for both recombinant protein and MAP. The *in vitro* method consists of incubating the purified recombinant protein product with purified MAP. The peptidase can be purified easily from hyperproducing *E. coli* strains, that express MAP at proportions of up to 20 per cent of total cell protein. The utility of these methods was demonstrated by the removal of methionine from interleukin-2 (Met-Ala-Pro-) and ricin A (Met-Ile-Phe-) both *in vivo* and *in vitro*. Recombinant proteins are normally exposed to the native activity of MAP in wild-type cells, but this low concentration often results in incomplete processing. Using an *E. coli* strain with extra MAP-bearing plasmids for the *in vivo* method increases the level of MAP in the cell, assuring more complete cleavage of the N-terminal methionine. In the *in vitro* method, proteins are incubated with a high concentration of enzyme, and the length of incubation can be adjusted to accomplish complete processing. □

Arie Ben-Bassat and Keith Bauer are at the Department of Fermentation R&D, Cetus Corporation, Emeryville, California 94608, USA. For more information fill in Reader Service No. 100.

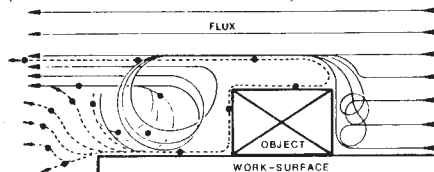
1. Sherman, F., Stewart, J.W. & Tsunasawa, S. *BioEssays* **3**, 27-31 (1985).
2. Flinta, C., Persson, B., Jornvall, H. & von Heijne, G. *Eur. J. Biochem.* **154**, 193-196 (1986).
3. Ben-Bassat, A. et al. *J. Bact.* **169**, 751-757 (1987).
4. Harris, T.S.R. *Genet. Engng* **4**, 127-185 (1983).
5. Kohr, W.J., Keck, R. & Harkins, R.W. *Analyt. Biochem.* **122**, 348-351 (1982).
6. Blumberg, S. & Ben-Meir, D. (inventors) Bio-Technology General Corp. (applicant), International Publication Number W086/01229 (World Intellectual Property Organisation 1986).
7. Prescott, J.M. & Wilkes, S.H. *Meth. Enzym.* **44**, 530-543 (1976).

A glimpse of Bio-Expo

Below is a sampling of some of the products to be on exhibit at the Bio-Expo conference held on 24–28 March in Paris, France.

PHARMACIA will be showing off its range of products for biotechnology in its booth at Bio-Expo (Reader Service No. 101). Not the least of Pharmacia's offerings will be the CellPfect Transfection Kit, for optimizing high efficiency **transfection of eukaryotic cells** with DNA. The kit contains three solutions; two optimized buffers for calcium phosphate transfection for stable expression, and one DEAE-Dextran solution for transient expression. The kit is supplied with a detailed protocol for both stable and transient expression procedures, and one kit contains enough reagents for 25–100 transfections, depending upon the method used and size of the Petri dish.

With the **sterile storage boxes** for laminar flow hoods developed by ESI-Flufrance, objects needed for sterile procedures may be kept near at hand, without interfering with the laminar flow and introducing the possibility of contamination (Reader Service No. 102). Instead of



ESI-Flufrance's storage boxes go with the flow.

keeping objects on the work area surface where they can interrupt the air flow, using ESI-Flufrance's side-mounted box allows researchers to keep the work area clear of extra utensils until they are needed. The stainless-steel storage box is equipped with an internal ultraviolet

radiation source, so its contents are kept sterile. The box has a glass door that opens directly inside the work area, at an angle designed so that the pivoting axis of the door is always parallel with the direction of the flow, for minimum flow disturbance. ESI-Flufrance has developed the sterile storage boxes for both horizontal and vertical laminar flow hoods. As an option, a second glass door may be added on one of the outer faces of the storage box to convert it into a sterilization chamber.

With the **solids probe** for mass spectrometry developed by Kratos Analytical, analyses can be performed under precise temperature control (Reader Service No. 103). The £7,000 (UK) probe's tip is composed of stainless steel and has the heater

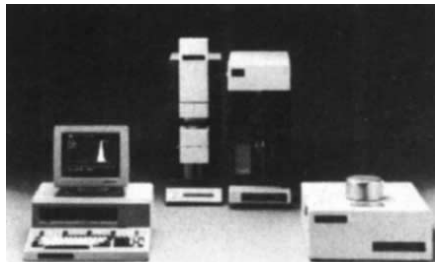


The Kratos Analytical probe and control unit.

connected directly to it, allowing rapid heat transfer from the sample, so the recorded and actual temperatures match exactly. According to Kratos Analytical, the solids probe tip temperature can exceed 800°C, and 600°C can be reached from ambient temperature in approximately one minute. The probe is also heatable in atmosphere without damage, and can be cooled, directly to the tip end, with liquid nitrogen or dry air.

Hewlett-Packard's new **analytical supplies catalogue** is out, and it now contains information on such products as sample preparation supplies, laboratory robot supplies, filters, vials, flow metres, and even a water purification system (Reader Service No. 104). The free 224-page catalogue also outlines HP's more traditional products: gas and liquid chromatography columns, syringes, chemical samples, and supplies for HP's chromatography and mass spectrometry instruments. The catalogue's appendices, including more than 100 capillary and liquid chromatography column applications, chromatographic equations, miscibility tables, conversion tables, and a glossary of chromatographic terms, make it a handy reference for analytical chemists.

Perkin-Elmer has recently introduced a new line of **thermal analysis** systems, the DELTA series (Reader Service No. 105). Presented as a lower-cost option for laboratories with limited budgets or those who are new to thermal analysis, the £18,500–45,000 (UK) systems are based on the Perkin-Elmer 3700 Data Station as instrument controller. According to Perkin-Elmer this system permits configuration of a low-cost thermal analysis system by providing complete control of a single



DELTA: the low-cost alternative for thermal analysis.

thermal analyser module. The library of software available for the DELTA series currently includes programs for instrument control, standard and advanced analyses, and multitasking options that allow the DELTA series to perform simultaneous data collection, analysis and plotting.

Flow is now distributing Gelaire **laminar air flow benches** and **Airone workstations** (Reader Service No. 106). The horizontal laminar air flow benches have automatic air speed control, and the range of five models offers up to 60 options. The benches are available as either table-top or freestanding models, with stainless steel or formica work surfaces. The vertical laminar air flow benches for tissue culture come in three sizes, and also sport automatic air speed control. The Airone fume containment workstations, also built by Gelaire, feature a wide range of activated carbon filters, making them a flexible system for the control of occupational hazards. The Airone range includes modular bench units and a mobile workstation.

Brownlee Labs, a division of Applied Biosystems, Inc., has a new product bulletin available describing the Brownlee SFC System One for **supercritical fluid chromatography** (SFC) (Reader Service No. 107). The free brochure contains a complete description of System One, plus pointers on when to use SFC, and the suitability of different sample types for analysis by SFC. Illustrative chromatograms are also included. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENT

SINGER INSTRUMENTS Singer Instruments Xenopus oocytes injection?



SINGER MKI MICROMANIPULATOR

We also offer a Microsyringe Kit.

Write to:

Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50